National Library of Canada
Bibliothèque nationale du Canada

Canadian Theses Service
Service des thèses canadiennes

Ottawa, Canada
K1A 0N4

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

NL-339 (r. 88/04) c

Canada

UNIVERSITY OF ALBERTA

ANALYSIS OF GENE FLOW IN *POPULUS* SPECIES USING FLAVONOIDS AND ISOENZYME MARKERS

BY

AGNES SYLVIE VANENDE

(C)

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH IN PARTIAL FULFILLMENT OF THE REQUIREMENT FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF FOREST SCIENCE

EDMONTON, ALBERTA

FALL 1990

National Library of Canada | Bibliothèque nationale du Canada

Canadian Theses Service | Service des thèses canadiennes

Ottawa, Canada
K1A 0N4

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-315-65027-3

UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR : AGNES SYLVIE VANENDE

TITLE OF THESIS : ANALYSIS OF GENE FLOW IN *POPULUS* SPECIES USING FLAVONOIDS AND ISOENZYME MARKERS

DEGREE : DOCTOR OF PHILOSOPHY

YEAR THIS DEGREE GRANTED : 1990

PERMISSION IS HEREBY GRANTED TO THE UNIVERSITY OF ALBERTA LIBRARY TO REPRODUCE SINGLE COPIES OF THIS THESIS AND TO LEND OR SELL SUCH COPIES FOR PRIVATE, SCHOLARLY OR SCIENTIFIC PURPOSES ONLY.

THE AUTHOR RESERVES OTHER PUBLICATION RIGHTS, AND NEITHER THE THESIS NOR EXTENSIVE EXTRACTS FROM IT MAY BE PRINTED OR OTHERWISE REPRODUCED WITHOUT THE AUTHOR'S WRITTEN PERMISSION.

.....

76 allée des Places

33470 GUJAN-MESTRAS

FRANCE

Date : August 9th, 1990

UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

THE UNDERSIGNED CERTIFY THAT THEY HAVE READ, AND RECOMMEND TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH FOR ACCEPTANCE, A THESIS ENTITLED :

ANALYSIS OF GENE FLOW IN *POPULUS* SPECIES USING FLAVONOIDS AND ISOENZYME MARKERS

SUBMITTED BY AGNES SYLVIE VANENDE

IN PARTIAL FULFILLEMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Dr. Bruce P. Dancik .....

Dr. Keith E. Denford .....

Dr. Stephen J. Titus .....

Dr. J. Dave Jobson .....

Dr. Burton V. Barnes .....
for R.V. Barnes

Date : June 26th, 1990

Dedication

When fires are set, and windows blind with snow
We shall remember, with a yearning pang,
How in the poplars the first robin sang.

Archibald Lampman
(The passing of spring)

A mes parents,
très affectueusement

Abstract

In southwestern Alberta, three species and subspecies of *Populus* section *Tacamahaca* are sympatric and have been reported to exhibit introgression :

- *P. angustifolia* James
- *P. balsamifera* L. subsp. *balsamifera* L.
- *P. balsamifera* L. subsp. *trichocarpa* (T.&G.) Brayshaw.

Interbreeding between taxa is reported to occur readily with putative hybrid swarms being observed in disturbed areas which favour their establishment. The genetic background of these putative hybrid swarms is complex. Morphologically it is possible to infer the parental species but difficult to determine the degree to which each parent contributes its characteristics.

The organization of these putative hybrid populations was investigated using flavonoid chemistry and isozyme gel electrophoresis as well as classical morphology. A sample of 366 trees, representative of 20 southwestern Albertan locations was studied.

To avoid intra-individual variability, morphological characters were studied using "early leaves". Satisfactory distinctions between *P. angustifolia*, *P. balsamifera* and *P. angustifolia* x *P. balsamifera* putative hybrids were shown, with *P. balsamifera* subspecies exhibiting continuous variation of leaf shape among individuals.

The study of pollen viability and flower morphology was

limited by poor flowering rates. However, the data obtained on the flowering individuals revealed probable occurence of hybridization between *P. angustifolia* and *P. balsamifera* in the populations sampled.

Electrophoresis of winter buds was limited by the presence of resin and did not give satisfactory separation of the 3 taxa. However, the 9 enzyme systems studied inferred high polyclonality. Root material exhibited intra-individual variability and was discarded.

Flavonoid chemistry was shown to be the most useful method and permitted satisfactory distinction between taxa. *P. balsamifera* flavonoid profile is characterized by the presence of flavones which are absent from *P. angustifolia* in which only flavonols are present. Distinction between *P. balsamifera* subspecies is more subtle : subsp. *balsamifera* has a simpler glycosylation pattern than subsp. *trichocarpa*. Two thirds of the individuals studied exhibited profiles intermediate between the 3 pure taxa profiles. They can be classified in 6 levels of intermediacy. Subsp. *balsamifera* was found more strongly represented than subsp. *trichocarpa* in southwestern Alberta.

The numerous intermediate forms between subsp. *balsamifera* and subsp. *trichocarpa* suggested a lack of efficient sterility barriers among the putative parent taxa, confirming the two-subspecies-status of *P. balsamifera* proposed by Brayshaw (1965).

Résumé

Dans le sud-ouest de l'Alberta, trois espèces et sous-espèces de *Populus* de la section *Tacamahaca* sont sympatriques et ont été signalées présenter de l'introgression :

- *P. angustifolia* James
- *P. balsamifera* L. subsp. *balsamifera* L.
- *P. balsamifera* L. subsp. *trichocarpa* (T.&G.) Brayshaw.

Les croisements entre ces taxons ont été signalés avoir lieu fréquemment et des groupes d'hybrides putatifs ont été observés dans les milieux instables qui favorisent leur établissement. Les antécédents génétiques de ces groupes d'hybrides putatifs sont complexes. Morphologiquement, il est possible de déduire les espèces parentales, mais il est difficile de déterminer à quel degré chaque parent contribue ses caractéristiques.

L'organisation de ces populations d'hybrides putatifs a été étudiée en utilisant les flavonoides, l'électrophorèse ainsi que les techniques classiques de morphologie. Un échantillon de 366 arbres représentant 20 stations du sud-ouest de l'Alberta a été étudié.

Pour éviter la variabilité intra-individuelle, les caractères morphologiques ont été étudiés sur des "feuilles précoces". Une distinction satisfaisante entre *P. angustifolia*, *P. balsamifera* et les hybrides putatifs *P. angustifolia* x *P. balsamifera* est démontrée tandis que les

sous-espèces de *P. balsamifera* présentent une variation continue de la forme foliaire entre individus.

L'étude de la viabilité du pollen et de la morphologie florale a été limitée par une incidence de floraison médiocre. Cependant, les résultats obtenus sur les individus ayant fleuri révèlent l'existence probable d'hybridation entre *P. angustifolia* et *P. balsamifera* dans les populations échantillonnées.

L'électrophorèse des bourgeons hivernaux a été limitée par la présence de résine et ne donne pas une séparation satisfaisante des trois taxons. Par contre, les 9 systèmes enzymatiques étudiés démontrent une importante pluri-clonalité. Le matériel racinaire présente une variabilité intra-individuelle et a été écarté.

Les flavonoides ont été démontrés la méthode la plus utile en permettant une distinction satisfaisante entre les taxons. Le profile flavonique de *P. balsamifera* est caractérisé par la présence de flavones qui sont absents de *P. angustifolia* chez qui seuls les flavonols sont présents. La distinction entre les sous-espèces de *P. balsamifera* est plus subtile : subsp. *balsamifera* possède un modèle de glycosylation plus simple que subsp. *trichocarpa*. Deux tiers des individus étudiés présentent des profils intermédiaires entre les profils des taxons purs. Ceux-ci peuvent être séparés entre 6 niveaux intermédiaires. Subsp. *balsamifera* est représentée de façon plus importante que subsp. *trichocarpa* dans le sud-ouest de l'Alberta.

Les nombreuses formes intermédiaires entre subsp. *balsamifera* et subsp. *trichocarpa* suggèrent un manque de barrières de stérilité efficaces entre les taxons parentaux putatifs, confirmant le statut de deux sous-espèces de *P. balsamifera* proposé par Brayshaw (1965b).

Acknowledgments

My appreciation and gratitude first extends to Dr. Bruce P. Dancik and Dr. Keith E. Denford, my degree program supervisors. They suggested the project, provided the facilities and generous financial assistance to carry it out.

I am also most grateful to Dr. Stephen J. Titus and Dr. J. Dave Jobson for their contributions as members of my dissertation committee. I am equally grateful to Dr. Burton V. Barnes (University of Michigan) for his role as an external examiner and his precious critical comments.

As members of my Ph.D. candidacy examination committee, I thank Dr. Walter Moser and Dr. Ken O. Higginbotham.

I would also like to thank the World University Service of Canada and the Government of Canada for providing me with a four year scholarship. This project was also supported by a two year grant from the Canadian Forestry Service of Canada. Their financial support is gratefully aknowledged.

I thank environment Canada and Alberta Parks and Recreation for permitting me to collect in National and Provincial Parks of Alberta.

My appreciation extends to Dr. Stewart B. Rood for his generous cooperation, giving me access to his *Populus* collection. I also thank Dr. Jordan Louviere and Michael Fox for their help with the delicate matter of correspondence analysis.

My special thanks to Zuzu Gadallah, Julianne Bottrill and Karin Thirlwell for their assistance in the field and the laboratory; Mary Aleksiuk for introducing me to electrophoresis; Rudi Kroon for guiding me in the greenhouse experiments; Bridj Kohli, Lakshmi Nandakumar, William Gruezo, Steve Downie and Michèle Fortin, for their frienship and support during these five years.
I also thank the office staffs of the Departments of Forest Science and Botany who put up with me, my incessant demands and my funny accent.

Finally, I thank most gratefully Dr. Robert Gorenflot (Université d'Orsay, France), the supervisor of my French graduate degree, who advised me to go explore the world.

Table of contents

Chapter	Page
I - Introduction	--1
II - Systematics of the genus *Populus*	--5
1 - The genus *Populus*	--6
2 - Section *Tacamahaca* Spach.	--9
3 - The american species of section *Tacamahaca*	-14
a - *Populus angustifolia* James, narrowleaf cottonwood	-15
b - *Populus balsamifera* L. subsp. *balsamifera* L., balsam poplar	-16
c - *Populus balsamifera* subsp. *trichocarpa* (T. & G) Brayshaw, black cottonwood	-20
d - Subspecies *balsamifera* and subspecies *trichocarpa* in Alberta	-23
III - The putative hybrid poplars of the section *Tacamahaca* in southwestern Alberta : material and methods	-27
1 - Literature review	-27
2 - Study area	-31
3 - List of selected locations	-34
4 - Additional locations	-40
5 - Methods : correspondence analysis	-42
a - Method presentation	-42
b - Correspondence analysis example	-45
c - Data structure considerations	-54
IV - Morphology	-57
1 - Definition	-57
2 - Literature review of the morphological studies on North American poplars	-57
3 - Material and technique	-59
a - Selection of material	-59
b - Collection and leaf selection	-59
c - Choice of characters	-62

4 - Data treatment -68
5 - Results -69
6 - Intra-individual variability of the morphological characters -75

V - Flower morphology and pollen study -79
1 - Flower morphology -79
2 - Pollen viability study -82

VI - Chemotaxonomy -88
1 - Definition and history -88
2 - Flavonoids -89
3 - Literature review of the flavonoids in genus *Populus* -92
4 - Material and technique -94
5 - The synthetic profile -96
6 - The species profiles -99
7 - Flavonoid pattern in southwestern Alberta 108
8 - Intra-individual variability of the flavonoid profile 118

VII - Allozyme electrophoresis 120
1 - Definition and theory 120
2 - Literature review of electrophoretic studies in *Populus* 122
3 - Material and technique 124
a - Selection of material 124
b - Collection of material and preparation 126
c - Electrophoresis technique 129
d - Zymogram patterns 135
4 - Results 135
a - Percentage distribution of each genotype 135
b - Correspondence analysis 139
c - Frequency distribution of each genotype in the locations 142
d - Synthesis 145
5 - Intra-individual variability of the allozyme pattern 146

VIII - General discussion and conclusions 148
1 - Method comparison 148
2 - Update of Brayshaw's distribution maps 153

Bibliography 156

Appendix 1 - Correspondence analysis program 170

Appendix 2 - Morphological data 176

Appendix 3 - Flavonoid data 181

Appendix 4 - Electrophoresis data 186

List of tables

	Page
Table 1 - Characteristics of *Populus* sections (Eckenwalder 1977, F.A.O. 1980).	--8
Table 2 - Species of the genus *Populus* (F.A.O. 1980, Dickman and Stuart 1983).	-10
Table 3 - Distinctive characters of *P. balsamifera* subsp. *balsamifera* and subsp. *trichocarpa*.	-24
Table 4 - Morphological characters of *P. angustifolia*, the two subspecies of *P. balsamifera* and intermediate specimens (Brayshaw 1965b).	-30
Table 5 - Altitude, latitude, longitude of the 21 *Populus* locations selected for the study of balsam poplars in southwestern Alberta.	-39
Table 6 - Correspondence analysis example : presence-absence table of 8 chemicals (02 to 33) in leaf extracts of 10 poplar trees (T10 to T65).	-46
Table 7 - Correspondence analysis example : presence-absence of 8 chemicals in 10 poplar trees. a - Inertias (eigenvalues), percentages of total inertia and cumulative percentages b - Principal coordinates for rows (trees) c - Principal coordinates for columns (chemicals).	-48

Table 8 - Correspondence analysis example : presence-absence of 8 chemicals in 10 poplar trees.
a - Absolute contributions of row and column points to inertia -51
b - Relative contributions of row and column points to inertia. -53

Table 9 - Morphological characters used in various studies of North American poplars. -65

Table 10 - Values of the morphological characters for the examples shown on figure 14; these values are the means calculated from the measurements done on six leaves per individual. -73

Table 11 - Morphological characters of *P. angustifolia*, the two subspecies of *P. balsamifera* and intermediate specimens, derived from the corespondence analysis results. -74

Table 12 - Morphological data (discrete) for both collections of Kananaskis 4 (1986 and 1987). -78

Table 13 - Pollen viability of the 54 *Populus* individuals tested; collection March 1989. -85

Table 14 - Literature review of the flavonoids in *Populus*. -93

Table 15 - Chromatographic characteristics of the flavonoid zones recorded on the synthetic profile of the *Populus* taxa studied. -98

Table 16 - Flavonoid structures isolated from the *Populus* taxa studied.
a - U.V. absorption maxima (nm). 100
b - Rf. and probable identity. 101

Table 17 - Presence-absence of the identified flavonoid structures in the flavonoid profiles of *Populus* taxa studied.
a - All structures. 105
b - Key flavonoid structures for the identification of the taxa studied. 106

Table 18 - Distribution of the different types of flavonoid profiles in the 21 *Populus* locations. 115

Table 19 - Literature review of electrophoretic studies in *Populus*. 123

Table 20 - Enzyme systems investigated in buds and roots of *Populus* taxa studied. 125

Table 21 - Extraction buffer, gel buffer and electrode buffer recipes used for the electrophoresis of *Populus* samples. 127

Table 22 - Enzyme stain recipes used for the electrophoresis of *Populus* samples. 131

Table 23 - Substrate solutions used in enzyme stain recipes for the electrophoresis of *Populus* samples. 134

Table 24 - Loci and alleles observed for the nine enzyme systems studied in the *Populus* samples. 138

Table 25 - Percentage distribution of each genotype in the population of *Populus* taxa studied. 140

Table 26 - Frequency distribution of each genotype in the 21 *Populus* locations (only for loci exhibiting variation). 143

List of figures

Page

Figure 1 - Natural range of the genus *Populus* (F.A.O. 1980). --7

Figure 2 - Natural range of the section *Tacamahaca* (F.A.O. 1980). -13

Figure 3 - Natural range of *Populus angustifolia* James (Little 1976). -17

Figure 4 - Natural range of *Populus balsamifera* L. subsp. *balsamifera* L. (Little 1971). -19

Figure 5 - Natural range of *Populus balsamifera* ubsp. *trichocarpa* (T. & G.) Brayshaw tle 1971). -22

Figure 6 - Range limits of *P. angustifolia*, *P. balsamifera* subsp. *balsamifera* and subsp. *trichocarpa* in southwestern Alberta (according to Brayshaw 1965b). -29

Figure 7 - Study area and 21 selected *P. angustifolia* and *P. balsamifera* locations in southwestern Alberta. -33

Figure 8 - Additional *P. balsamifera* locations. -41

Figure 9 - Correspondence analysis example : presence-absence of 8 chemicals in 10 poplar trees, first 2 principal axes plot. -49

Figure 10 - Dimension and shape variability of the "late leaves" (#6 to 15) of *P. balsamifera* on a one-year-old twig (Vanende 1984). -60

Figure 11a - Silhouettes of *Populus balsamifera* leaf samples from High River location (HR-03) and Daisy May location (DM-04). -63

Figure 11b - Silhouettes of *Populus angustifolia* leaf samples from Stand Off location (SO). -64

Figure 12 - Quantitative characters measured on the poplar leaves in this study. -66

Figure 13 - Correspondence analysis plot (first 2 principal axes) of the transformed morphological data matrix. -70

Figure 14 - Correspondence analysis plot (first 2 principal axes) of the transformed morphological data matrix; the different trends of leaf shape exhibited in the population are illustrated with leaf examples. -72

Figure 15 - Flavonoid 15 carbon unit. -91

Figure 16 - Flavonoid classes and interrelationships (Ribéreau-Gayon 1968, Markham 1982). -91

Figure 17 - Synthetic flavonoid profile of the *Populus* taxa studied; each flavonoid zone is delimited and numbered. -97

Figure 18 - Identified flavonoid structures of the synthetic flavonoid profile of *Populus* taxa studied. 102

Figure 19 - Flavonoid profiles of the pure *Populus* taxa studied :
a - *P. angustifolia* from Gunnison, Colorado, and Stand-Off, Alberta
b - *P. balsamifera* subsp. *balsamifera* from Plamondon and High Level, Alberta
c - *P. balsamifera* subsp. *trichocarpa* from Boston Bar and Savona, British Columbia. 104

Figure 20 - Correspondence analysis plot (first 2 principal axes) of the flavonoid data matrix. 109

Figure 21a - Intermediate flavonoid profiles observed in balsam poplars of southwestern Alberta : 3 types of intermediate profiles between *P. angustifolia* and *P. balsamifera*. 112

Figure 21b - Intermediate flavonoid profiles observed in balsam poplars of southwestern Alberta :
1 - *P. balsamifera* intermediate
2 - *P. balsamifera* intermediate tendency subsp. *balsamifera*
3 - *P. balsamifera* intermediate tendency subsp. *trichocarpa*. 113

Figure 22 - Correspondence analysis plot (first 2 principal axes) of the flavonoid profile distribution in the 21 *Populus* locations. 116

Figure 23 - Zymogram patterns observed in the *Populus* taxa studied (winter buds). 136

Figure 24 - Correspondence analysis plot (first 2 principal axes) of the electrophoresis data matrix. 141

Figure 25 - Correspondence analysis plot (first 2 principal axes) of the frequency distribution data matrix (frequency of each genotype in the 21 *Populus* locations). 144

I - Introduction

The genus *Populus* occurs throughout most of the forests of the temperate and cold regions of the northern hemisphere.

Uniform populations can develop naturally where there is no serious competition with other major tree species as components of the forest. However, poplars are "light demanding" species and normally develop on open sites behaving as pioneer occupants of disturbed spaces. They appear also on sites unused by agriculture such as river banks and channels and alluvial areas, all poor in humus and unable to sustain a worthwhile agricultural crop. The colonization ability of *Populus* makes it very successful in these areas. Because of the ease of propagation by cuttings, many pioneers planted them near their homes, around fields, along roads and ditches, either to furnish wood, shade or windbreaks. In some areas, poplars have been associated with agriculture since antiquity, as in the Near East where they are often the only trees that will grow. For such reasons, they are called the trees of the countries with no forests. This long association is revealed by the name of the genus. *Populus* comes from "populus", meaning people in Latin and signifying a tree grown around public places. The word poplar is one of the few tree names to be shared in different forms by several modern European tongues such as French, Dutch, German, Danish, Finish, English, Welsh, and

Italian (Dickman and Stuart 1893).

The primitive way of growing poplars by just thrusting a shoot into the soil soon evolved into a legitimate method of silviculture. At the beginning of the 18th century, *Populus deltoides* a North American species was introduced to Europe, eventually resulting in the appearance of natural hybrids between European and American poplars of section *Aigeiros*. Since then, selection of new hybrids, development of new techniques and protection against pests and diseases have been a constant process.

The poplar is often proposed as the archetype of the forest tree manipulated by man in his search to control nature (Pauley 1949, Morhdiek 1983). After such domestication, it is difficult to delimit and describe the different species of the genus and their natural distributions. The systematics of the genus *Populus*, relatively simple at the level of section, and even at species level, becomes complex and often uncertain at the lower levels. Because of the existence of many hybrid forms, it is most often a fascinating but frustrating study of classification. A good example of this can be observed in southwestern Alberta where four different species and subspecies are sympatric and have been reported to hybridize and exhibit introgression : *Populus balsamifera* L. subsp.

balsamifera L., *P. balsamifera* L. subsp. *trichocarpa* (T. & G.) Brayshaw, *P. angustifolia* James and *P. deltoides* Marshall. Here, the four taxa have been reported to interbreed freely and putative hybrid swarms have been observed in disturbed areas which may favour the establishment of hybrids (Brayshaw 1965b, Rood 1985).

The variation present in the poplar populations of southwestern Alberta maybe due to a combination of undefined within-taxon variation, poor differentiation between taxa and/or possible hybridization among taxa. In fact, there are three distinct taxonomical problems in this area.

- Firstly, the genetic background of these putative hybrid swarms may be complex. Morphologically, it is possible to infer the parental species but almost impossible to determine the degree to which each parent contributes its characteristics.
- The second problem is the status of *Populus balsamifera*. Subspecies *balsamifera* and *trichocarpa* were considered as two distinct species until Brayshaw (1965b) suggested to change their status. However, this modification is not unanimously accepted (Eckenwalder 1977).
- Thirdly, the distinction itself of the two subspecies is difficult, as morphological characters are not reliable (Brayshaw 1965b).

The first objective of this study is to solve these three problems by using modern technologies such as flavonoid

chemistry and electrophoresis, as well as classical morphological approaches.

Brayshaw (1965b) delimited the distribution of *Populus angustifolia*, *P. balsamifera* subsp. *balsamifera* and *P. balsamifera* subsp. *trichocarpa* in southwestern Alberta. This area has been greatly modified by progressive damming of the river systems and subsequent extension of agriculture. The second objective of this study is to investigate possible changes (size, mortality, evolution, etc) of the poplar populations in southwestern Alberta since Brayshaw's study and to update his distribution maps.

II - Systematics of the genus Populus

The genus *Populus* belongs to the family *Salicaceae* along with three other genera, *Salix*, *Chosenia* and *Toisusu*. *Salicaceae* forms part of the order *Salicales* of the group *Amentiflorae* characterized by unisexual flowers with absent or insignificant perianths.

Male and female flowers in the *Salicaceae* are simple and mostly reduced to stamens or pistil, arranged in a more or less elongated catkin. The fruit is a dehiscent capsule, opening by two or more valves, with seeds being numerous and furnished with tufts of silky hairs commonly called "cotton".

The family has a number of characteristics which differ from most other wood-producing tree species :

1. - the *Salicaceae* reproduce freely by vegetative processes and not regularly by seeds
2. - the *Salicaceae* are dioecious
3. - hybridization is frequent, naturally or by controlled crossing (Muhle Larsen 1970, Zsuffa 1375).

1 - The genus *Populus*

Species of the genus *Populus* are widely distributed throughout the northern hemisphere, chiefly in the temperate zone (figure 1). Only *Populus euphratica* Oliv. in Central Africa is native to the southern hemisphere.

The genus *Populus* is divided into five sections : *Turanga*, *Leuce*, *Aigeiros*, *Tacamahaca* and *Leucoides*, each with particular distinguishing features (table 1).

The systematics of the genus *Populus* is characterized by much synonymy and misapplied names. Since the first description of the genus by Linnaeus (1753), many authors have tried to construct a satisfactory infrageneric classification. Over the years, sections, subsections and species have been reclassified, redivided and renamed. European and American authors frequently disagree on the latin name of certain species. For example, *Populus deltoides* Bartram ex Marshall has been ascribed 10 different names plus many varieties and subspecies, always differing somewhat in description.

The situation is complicated by natural and artificial hybridization in the genus, resulting in confusion and many nomenclatural problems. A taxonomic history of the genus *Populus* was published by Eckenwalder (1977).

Currently, most authors agree on the five sections named above with total species varying between 30 and 45 (Britton 1923, Rehder 1951, Hitchcock et al. 1964, Dickman and Stuart

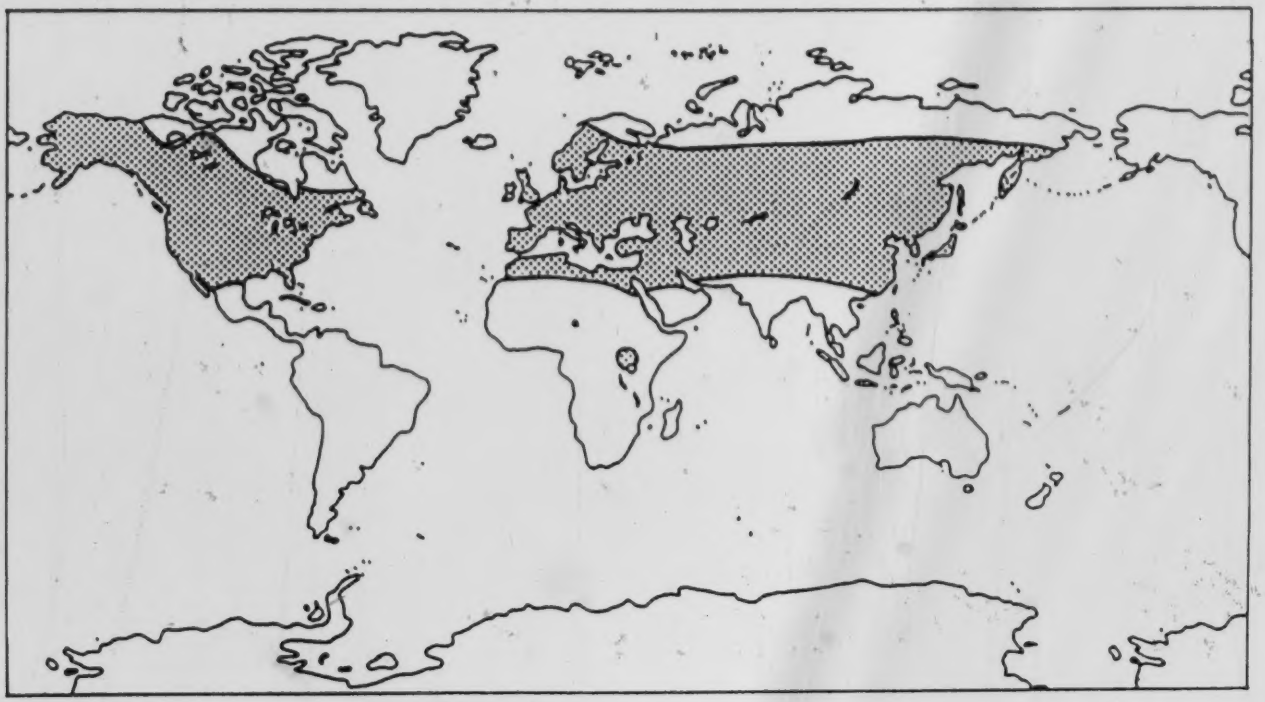


Figure 1 - Natural range of the genus *Populus* (F.A.O. 1980).

	Turanga	*Leuce*		*Leucoides*	*Tacamahaca*	*Aigeiros*
		Trepidae	*Albidae*			
Number of stamens	25-30	5-20	6-10	20-30	15-50	20-80
Number of carpels	3, rarely 2	2		2 or 3	2 to 4	2 to 4
Capsule shape	ovoid	narrow ovoid		ovoid	ovoid to subglobose	ovoid to subglobose
Leaves	heterophyllous	heterophyllous		not heterophyllous	heterophyllous	heterophyllous
Buds	small, downy	small, downy		slightly resinous	large, very resinous	large, very resinous
Geographical distribution	East and Central Asia, North Africa	T : North and mountainous regions of Eurasia and North America A : Circum-mediterranean, mountainous regions in East Europe		South East North America and Central Asia	North and mountainous regions of North America and Central Asia	North America and Mediterranean region

Table 1 - Characteristics of *Populus* sections (Eckenwalder 1977, F.A.O. 1980).

1983).

Thirty species are generally recognized (table 2). This list takes into account the most recent revision of the genus by Eckenwalder (1977), in which the addition of a sixth section, *Abaso*, is proposed. Section *Abaso* regroups species formerly included in section *Aigeiros* but differing in floral and vegetative characteristics. Eckenwalder (1977) also reported the presence in Mexico of two endemic species of the section *Leuce* sub-section *Albidae* : *P. simaroa* Rzedowski and *P. monticola* T. Brand.

Ten of the species listed in table 2 are native to North America and seven are utilized by the forest industry (Fowells 1965).

2 - Section *Tacamahaca* Spach.

Tacamahaca, or the balsam poplars, is the largest section in the genus. Balsam poplars take their name from the fragrant resin abundantly present in the winter buds of all species in this section.

Balsam poplars occur in northern latitudes in both Asia and North America (figure 2). In Asia, they occupy a wide area extending from the Arctic Circle to the southern flanks of the Himalayas. Their economic importance is poorly known, except for a few clones exported to North America and Europe, which have been used extensively in breeding

table 2 - Species of the genus *Populus* (F.A.O. 1980, Dickman and Stuart 1983).
* Native species to North America
** Native species to North America used in Forestry
+ Revision by Eckenwalder (1977)

Section	Species	Common name	Distribution
Turanga Bge.	-*P. euphratica* Oliv.	Euphrates poplar	S. Europe, W. and Cent. Asia, N. Africa, Kenya
Leucoides Spach.	-*P. lasiocarpa* Oliv.		China, only monoecious species of *Populus*
	-*P. ciliata* Wall	Himalayan poplar	Himalayan mountains
	-*P. heterophylla* L. (**)	Swamp cottonwood	S. W. United States
	-*P. wilsonii* Schneid.	Wilson poplar	China
Leuce Duby Subsection *Albidae*	-*P. alba* L.	white poplar	N. Africa, S. Europe, Cent. Asia
	-*P. monticola* T. brand. (+*)		endemic Mexico
	-*P. simaroa* Rzedowski (+*)		endemic Mexico
Subsection *Trepidae*	-*P. tremula* L.	European aspen	Most Europe, N. Africa, N. Asia
	-*P. adenopoda* Maxim.	Chinese aspen	Cent. and W. China
	-*P. davidiana* Schneid.	Korean aspen	Korea

Table 2 - continued.

	-*P. sieboldii* Miq.	Japanese aspen	Japan
	-*P. grandidentata* Michx. (**)	Largetooth aspen	N. E. U.S.A., adjacent Canada
	-*P. tremuloides* Michx. (**)	Trembling aspen	N. America except Central Plains and S. E. part
Tacamahaca Spach.	-*P. laurifolia* Ledeb.	Laurel Poplar	Siberia
	-*P. szechuanica* Schneid.		China
	-*P. koreana* Rehd.	Korean poplar	Korea
	-*P. maximowiczii* Henry	Japanese poplar	Japan
	-*P.simonii* Carr.	Simon poplar	China, Korea
	-*P. yunnanensis* Dode		China
	-*P. suaveolens* Fisch.		Turkestan to Siberia
	-*P. cathayana* Rehd.		China, Korea
	-*P. tristis* Fisch.	Himalayan balsam poplar	Cent. Asia
	-*P. angustifolia* James (*)	Narrowleaf cottonwood	Rocky Moutains and Central Plains in U.S.A. and Canada

Table 2 - continued.

	-*P. balsamifera* L. subsp. *balsamifera* L. (**)	Balsam poplar	N. of North America
	subsp. *trichocarpa* (T. & G.) Brayshaw (**)	Black cottonwood	N. W. America
Aigeiros Duby	-*P. deltoides* Bart. ex Marsh. (**)	Eastern cottonwood	E. of U.S.A. and Canada
	-*F. fremontii* Wats. (**)	Fremont cottonwood	S. W. of U.S.A.
	-*P. nigra* L.	Black poplar	Eurasia
Abaso Ecken. (+)	-*P. mexicana* Wesm. (*)	Mexican poplar	Mexico

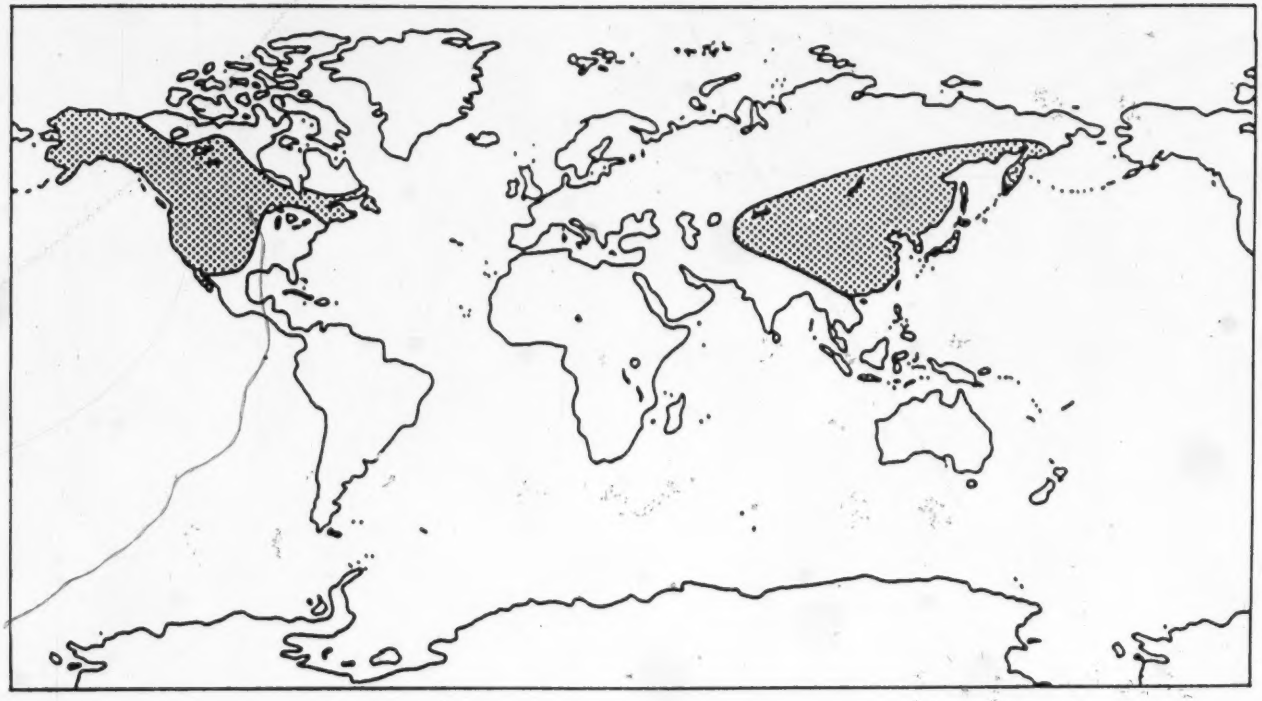


Figure 2 - Natural range of the section *Tacamahaca* (F.A.O. 1980).

programs (Stout and Shreiner 1933). In North America, they are present from the Arctic Circle across Canada, Alaska and the northern United States. They are also present on the Pacific Coast to northern Mexico, where they are restricted to the cordillera. Natural populations have great economic value, and several clones are planted extensively as ornamentals and windbreaks (F.A.O. 1980).

3 - The American species of section *Tacamahaca*

Three *Tacamahaca* taxa are native to North America :

1. - *Populus angustifolia* James
2. - *Populus balsamifera* L. subsp. *balsamifera* L.
3. - *Populus balsamifera* L. subsp. *trichocarpa* (T. & G.) Brayshaw

As in many other species of the genus *Populus*, these three taxa exhibit heterophylly. On the same tree, two types of foliage are regularly produced : early leaves and late leaves (Critchfield 1960). Two different kinds of shoot are also present, short shoots a few mm long bear only early leaves, long shoots (heterophyllous shoots) bear both a set of early leaves and a set of late leaves. The early leaves are produced early in the spring from embryonic leaves present in the winter bud, while late leaves are produced throughout the growing season accompanied by stem elongation.

Early leaves and late leaves not only differ in shape, but early leaves are generally smaller, narrower and thinner than late leaves and also have proportionally longer petioles and marginal teeth mostly devoid of glands (Eckenwalder 1977).

The leaves of the American balsam poplars are generally lanceolate to ovate with a terete petiole, channeled above and shorter than the blade. Winter buds are elongate, sharp-pointed and glabrous.

Typically pioneer species, American balsam poplars are trees of riparian habitats in boreal and montane regions.

a - *Populus angustifolia* James, narrowleaf cottonwood

This species was first described by James in 1823. The smallest of the three taxa of the section *Tacamahaca* in Alberta, this tree is usually 5 to 10 m tall. Its main characteristics include a slender trunk with a light yellowish grey bark and long narrow leaves. The crown is narrow pyramidal with ascending branches and glabrous foliage.

The leaves are light yellowish green above, paler or white beneath and often resin stained. Their size is small, 5 to 7.5 cm long and with the width less than a third of the length. They are lanceolate, ovate-lanceolate or ovate, with a margin finely toothed to the tip. Early leaves are smaller than the late leaves and generally acute to obtuse. Late leaves are acuminate.

The petiole is very short, less than a third of leaf length and glandless, semi-circular in section and flattened near the base of the blade (terete).

The fruit is an ovate, glabrous, 2-valved capsule with a short pedicel.

Populus angustifolia is distributed in the Rocky Mountains and Plains from southern Alberta to the Mexican border, mainly along stream banks (figure 3). It can be found up to 1600-2400 m. In Alberta, it is restricted to stream banks of the southwest corner of the Province. *Populus angustifolia* is a pioneer species, one of the first plants to become established on gravel and sand bars newly exposed in disturbed areas. Its intolerance to shade combined with its small size make it unable to compete with other species, especially other poplars. It is generally absent from mature stands with a closed canopy. This species is especially scarce in its typical form, where it has been reported to be outnumbered by taller hybrids with other poplars in most populations (Hosie 1961, Brayshaw 1965b).

b - *Populus balsamifera* L. subsp. *balsamifera* L., balsam poplar

This subspecies was first described as a species by Linnaeus in 1753. A tall tree, 20 to 30 m high, subspecies *balsamifera* has a long straight cylindrical trunk with a narrow, open crown of ascending branches. The bark is greenish and smooth, turning gray and scaly in older trees.

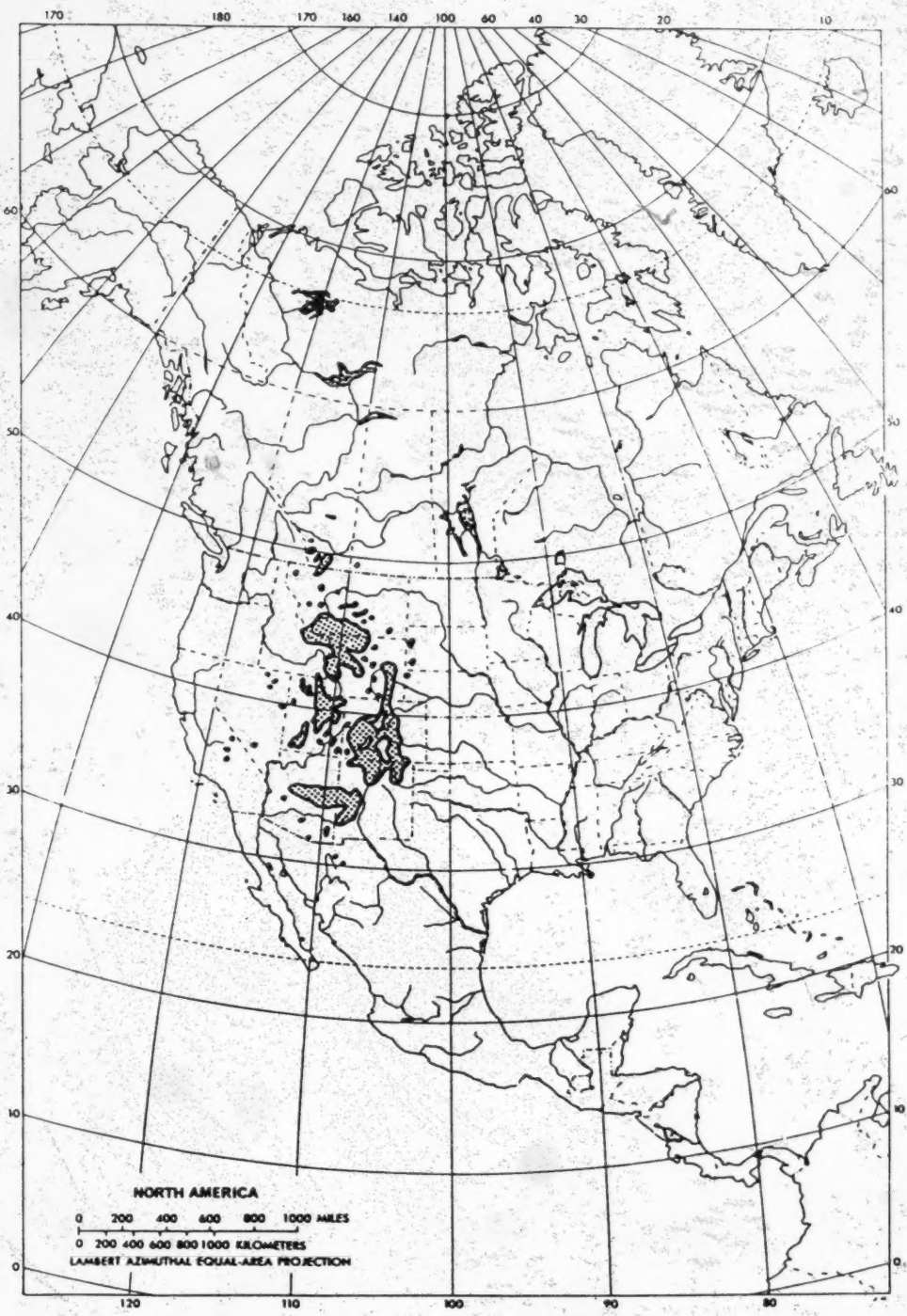


Figure 3 - Natural range of *Populus angustifolia* James (Little 1976).

The early leaves are 7.5 to 10 cm long and 4 to 9.5 cm wide. The upper surface is dark green, while the lower surface is glaucous and resin stained. The thickness of the leaves gives them a leathery texture. The early leaves are ovate to lanceolate, rounded at the base with an obtuse apex. Late leaves are ovate, rounded or cordiform at their bases and acuminate; they are generally larger than the early leaves. The petiole is 3 to 5 cm long, terete, channeled above, glabrous or puberulent. There is often a pair of glands at the junction with the blade.

The male flowers have 12 to 20 stamens with pinkish anthers. The fruit is an ovate, glabrous, 2-valved capsule with a short pedicel.

Two varieties have been distinguished :

1. - var. *subcordata* Hylander, Balm of Gilead, which has heart-shaped leaves, hairy on the lower surface
2. - var. *candicans* A. Gray, the natural origin of this variety is uncertain, as it seems to be escaped from cultivation (Britton 1923, Hosie 1961).

Subspecies *balsamifera* is distributed throughout the boreal forest from Newfoundland and Labrador extending westward across Canada to interior Alaska, up to the continental divide. It is also present in the Great Lakes-St Lawrence and Acadian forest regions (figure 4).

A pioneer species, very successful on river alluvium and

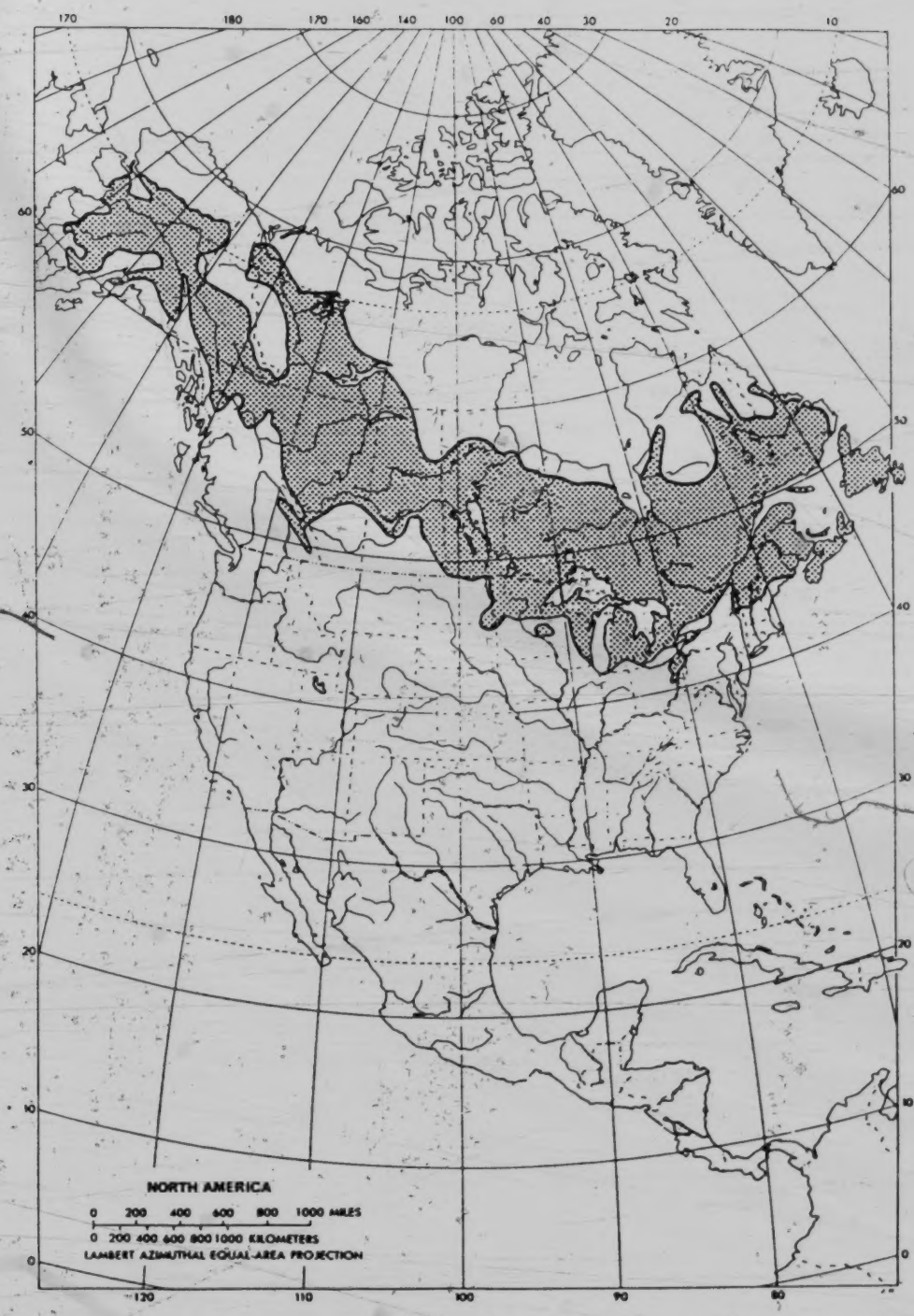


Figure 4 - Natural range of *Populus balsamifera* L. subsp. *balsamifera* L. (Little 1971).

glacial moraines, it is also an early successionnal species following fire and grows in the Boreal forest along the northern limit of tree growth. It is represented by pure populations or mixed with willow, alder, spruce, birch, etc. In southwestern Alberta, subspecies *balsamifera* is restricted to river banks.

c - *Populus balsamifera* L. subsp. *trichocarpa* (T. & G.) Brayshaw, black cottonwood

This subspecies was first described as a species by Torrey and Gray in 1843. The largest of the American balsam poplars, subspecies *trichocarpa* can reach 40 m high in the forest of the Pacific Northwest where it is also the tallest hardwood tree. In Alberta, it reaches 20 m in height.

The trunk is straight and cylindrical with a narrow columnar crown of ascending branches, and bark deeply furrowed in older trees.

The early leaves are 7 to 12 cm long and 4.5 to 12 cm wide, dark green above, glaucous white and resin stained below with a leathery aspect. Leaf shape varies considerably, from deltoid-ovate in the Pacific Northwest to heart shaped in California and lanceolate in the central eastern part of its range (Idaho, Montana, Central Oregon). Late leaves are frequently cordate at the base and early leaves mostly ovate.

The petiole is pubescent, 2 to 6 cm long, very similar in shape to subspecies *balsamifera*.

Male flowers have 40 to 60 stamens with reddish purple anthers.

The fruit is a subglobose to broadly ovate 3-valved capsule, usually pubescent.

A variety has been distinguished : var. *hastata* Dode which occurs throughout the range of subspecies *trichocarpa* and is characterized by its narrow and very acuminate leaves, and glabrous capsules.

Subspecies *trichocarpa* can be considered as the western counterpart of subspecies *balsamifera*. Its range extends along the Pacific coast from Alaska (Kodiak Island) to the Baja California. Not restricted to the coast, it is also present in the Great Basin (Nevada, Utah) and in Idaho and Central Montana (figure 5).

As with other *Populus* taxa, subspecies *trichocarpa* is a pioneer species present on riverbanks, alluvial flats and glacial moraines. Dense populations are frequent along rivers, but it also forms an important element of the Pacific Northwest forest, mixed with Douglas-fir, western red cedar, white spruce, etc. It occurs from sea level to high elevations (1500-2100 m) in the northern part of its range. In the southern part, it is most abundant from 1000 m to 1800 m, but can grow successfully up to 2700 m in California.

Figure 5 - Natural range of *Populus balsamifera* L. subsp. *trichocarpa* (T. & G.) Brayshaw (Little 1971).

d - Subspecies *balsamifera* and subspecies *trichocarpa* in Alberta

The identification of these two closely related taxa has always been a problem for taxonomists. Until 1965, when Brayshaw revised the genus *Populus* in Alberta, they were considered as two distinct species :

1. - *Populus balsamifera* L.
2. - *Populus trichocarpa* Torrey & Gray

The characters used to separate the two subspecies are summarized in table 3. Vegetative characters are not totally satisfactory, most are subjective and dimensions overlap almost completely.

Flower and fruit characters seem to be the only ones to be reliable in differentiating the two taxa (Brayshaw 1965a). Unfortunately, flowering rates and dates are highly variable from one year to the next. Furthermore, male flowers have a very short life which makes their collection difficult. For this reason, most herbarium specimens are without male flowers.

For many years, taxonomists used the locality of collections as an aid to identify herbarium specimens. Specimens from West of the Rocky Mountains were assigned to *Populus trichocarpa* and specimens East of the Rocky Mountains were assigned to *Populus balsamifera*.

In Alberta, British Columbia and Alaska, the precise geographic boundaries of the two subspecies are unclear. Where the ranges overlap, hybridization has been reported to

Table 3 - Distinctive characters of *P. balsamifera* subsp. *balsamifera* and subsp. *trichocarpa*.

	subspecies *balsamifera*	subspecies *trichocarpa*
Crown	narrow open crown, somewhat columnar	narrow open crown
Bark (on older trees)	scaly	deeply furrowed
Early leaves -width/length ratio	0.5-0.95	0.6-1.0
-base	mostly ovate	ovate, cordate or subcordate
Petiole	usually glabrous, sometimes pubescent	pubescent
Male flowers	12 to 20 stamens pinkish anthers	40 to 60 stamens purple anthers
Fruit	ovate, glabrous 2-valved capsule	subglobose to broadly ovate, pubescent 3-valved capsule

be frequent and identification almost impossible (Brayshaw 1965a 1965b, Vierek and Foote 1970, Rood 1985, Rood et al. 1986). In fact, most authors who studied non-sterile specimens studied only female flowers (Brayshaw 1965a, Vierek and Foote 1970). Since poplars are dioecious, this allows for the identification of approximatively half of the population.

Brayshaw (1965a) studied populations of the two subspecies in Alberta and British Columbia using the female flower characters. Most of the specimens examined proved to be intermediate, with 2-valved and 3-valved capsules on the same tree. Brayshaw concluded that in cordilleran and adjacent regions, the two taxa should be considered as a single species with two subspecies :

1. - *Populus balsamifera* L. subsp. *balsamifera* L.
2. - *Populus balsamifera* 1. subsp. *trichocarpa* Torrey and Gray.

Vierek and Foote (1970) came to the same conclusion when studying putative hybrid populations in Alaska. Both authors proposed the same hypothesis to explain this situation. During the Pleistocene, the advances of the ice sheets divided the common ancestral stock of the species into an eastern population and a western population. While isolated, these two populations evolved gradually into the two present subspecies. When the ice sheets began to retreat, both taxa expanded rapidly as pioneers to recolonize the ice-free lands. The isolation was not long enough to establish

effective sterility barriers between the two taxa and once in contact, hybridization and introgression occurred, to produce the actual mixed population. This contact, however, was limited by the presence of the Rocky Mountains. This new status has generally been accepted in Alberta (Packer 1983, Rood 1985). Hultèn (1968a, 1968b) extended this new status to the whole range of *Populus balsamifera* because of the apparent lack of sterility barriers between them. Eckenwalder (1977) reported a similar situation for the two species of the section *Aigeiros*, *P. fremontii* and *P. deltoides*, which exhibit the same degree of divergence in the central Plains of U.S.A.

III - The putative hybrid poplars of the section Tacamahaca in southwestern Alberta : material and methods

1 - Literature review

Brayshaw (1965b) identified an area of hybridization including the plains of southern Alberta from the Red Deer river southward. The presence of hybrids between subspecies *balsamifera* and subspecies *trichocarpa* was reported to occur northward up to the Peace river and also in British Columbia, but putative trispecific hybrid swarms also involving *Populus angustifolia* were limited to southern Alberta.

This area corresponds to the Northern Prairie defined by Moss (1932), extending on dark brown soil. It is limited in the North by Parkland, to the South and East by Southern Prairie and to the West by the Cordilleran Forest.

The prairies are generally treeless and covered by grassland, so poplars are not as dominant as they are in the boreal forest or even in Parkland. In the Prairies, poplar populations represent outliers and are mostly limited to riverbanks, being present along the rivers which flow in coulees deep enough to provide them with shelter from the wind. Along these rivers, poplars form gallery forests, accompanied by species of willows and birches.

The three taxa of the section *Tacamahaca* reach their range

limits in this area (figure 6). Subspecies *balsamifera*, a boreal forest tree, is present from the Bow River northward; subspecies *trichocarpa*, the western taxon, enters Alberta following the streams along the foothills; *Populus angustifolia*, a southern species, is represented by outlier populations along streams near the foothill zone.

The capacity for growth and reproduction of these trees is limited by the extreme climatic conditions of the prairies. As a consequence, the establishment of hybrids may be easier in this area because of lower competition from parental stocks. Also, some of the hybrids might be better fit to stand cold, drought, and wind exposure than their parents. Riverbanks colonized by poplars are very unstable. A high erosion rate, frequent floodings and deposition of sediments favor the movement of river channels. New open areas are constantly being formed, which being free of competition, might be favorable to the establishment of hybrids and their progeny (Grant 1971). In most of the area, the distinction between subspecies *balsamifera* and subspecies *trichocarpa* was impossible and only a common range limit (*P. balsamifera* intermediates) was recorded by Brayshaw (1965b).

Within the reported hybridization zone, putative hybrid individuals possess morphological characters with all degrees of intermediacy between the pure species (table 4). The epithet of *Populus x Brayshawii* was given to these hybrids (Packer 1983). Rood (1985) observed that because of the complexity of this continuous variation, it is

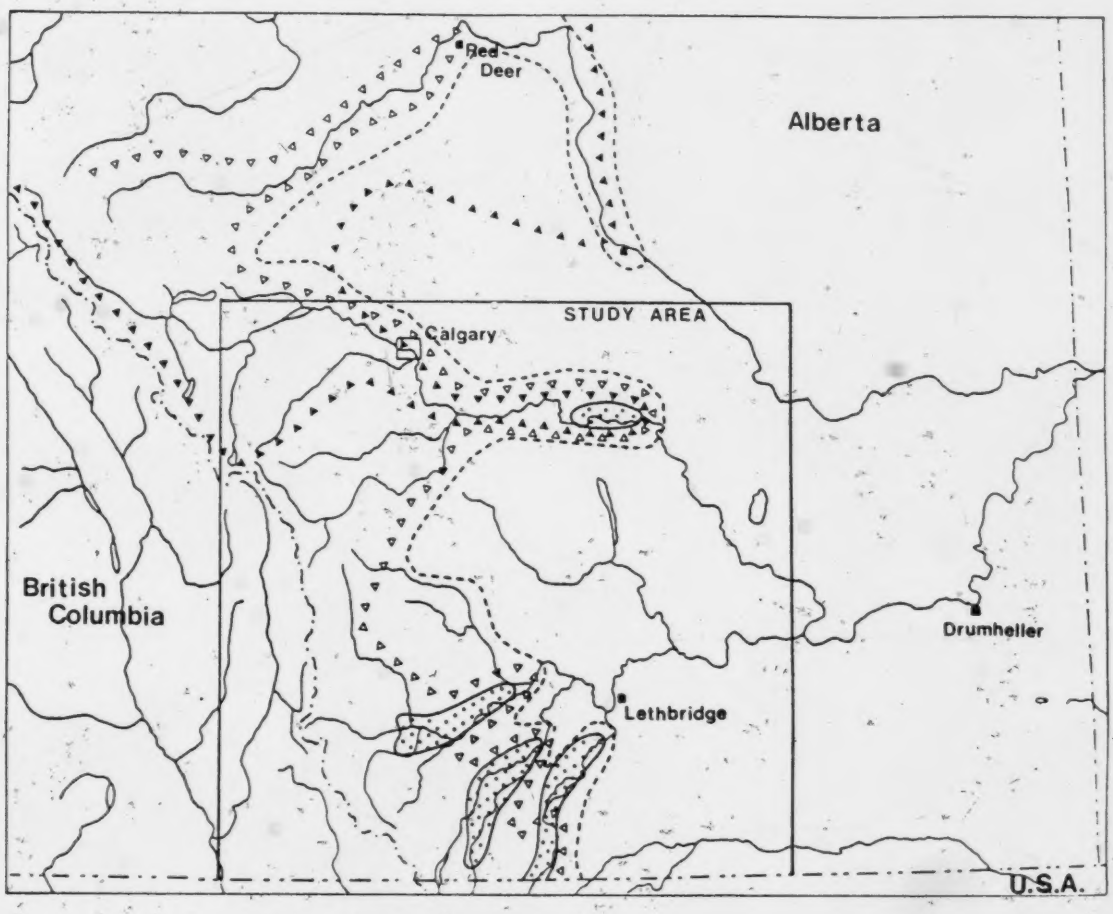


Figure 6 - Range limits of *P. angustifolia*, *P. balsamifera* subsp. *balsamifera* and subsp. *trichocarpa* in southwestern Alberta (according to Brayshaw 1965b).

P. angustifolia

P. balsamifera subsp. balsamifera

P. balsamifera subsp. trichocarpa

P. balsamifera indeterminates ----

Table 4 - Morphological characters of *P. angustifolia*, the two subspecies of *P. balsamifera* and intermediate specimens (Brayshaw 1965b).

	P. angustifolia	Intermediate specimens	P. balsamifera (2 subspecies)
Leaves			
- W/L	0.15-0.30	0.30-0.60	0.50-1.0
- apex	acute or acuminate	acute or acuminate	acuminate
- base	cuneate	rounded or cuneate	rounded to cordate
Petiole			
- P/L	<0.33	0.33-0.40	0.33-0.66
- cross section	semi-circular	lunate to terete	terete

preferable to determine the parent components present in a putative hybrid rather than give a different name to all hybrid combinations.

Brayshaw (1965b) indicated that this continuous variation was the consequence of many generations of interbreeding resulting in populations with a very complex genetic background. This continuous intergradation is only limited by linkage and natural selection, for example, narrow leaves never have long petioles, and wide leaves never have short petioles.

Another result of the long interbreeding may be the progressive disappearance of *Populus angustifolia* from Alberta where it is outnumbered by its putative hybrids. The species in its typical form is rare in Alberta, and most individuals exhibit traits common to other poplar species.

2 - Study area

To study the putative hybrid swarms of southern Alberta, the list of collection sites of Brayshaw (1965b) was followed initially and then modified to fit the objectives of the present study.

Brayshaw studied all the species of *Populus* present in Alberta. In the present study, *P. deltoides* section *Aigeiros* and *P. tremuloides* section *Leuce* are not included.

P. deltoides has been reported to hybridize with the species

of the *Tacamahaca* section :

1. - with *P. angustifolia*, the hybrids are named *P. x acuminata* Rydb.
2. - with the two subspecies of *P. balsamifera*, the hybrids are named *P. x jackii* Sarg.

These two groups of putative hybrids are fairly rare in southwestern Alberta and generally found separate from the main hybrid swarms.

P. tremuloides is not involved in the putative hybrid swarms as it does not hybridize with species of the *Tacamahaca* section.

Brayshaw's study area was also exclusively limited to the prairie. In the present study, it seemed necessary to extend the area in the foothills to check if poplar populations there were affected by hybridization.

The study area is limited to the north by the Bow River, to the south by the U.S.A. border and to the west by the Rocky Mountains. The eastern limit is in fact the balsam poplar's limit itself. The study was extended east as far as balsam poplar stands could be found, which is just east of Lethbridge. A very small population in Taber was included in the collection list, but should be considered as an outlier. A total of 21 locations were selected in this area in order to represent most of the river systems of southwestern Alberta (figure 7). However, no balsam poplar populations were found on the Little Bow River. The coulee of this river is too shallow to support poplar stands (Brayshaw 1965b). In

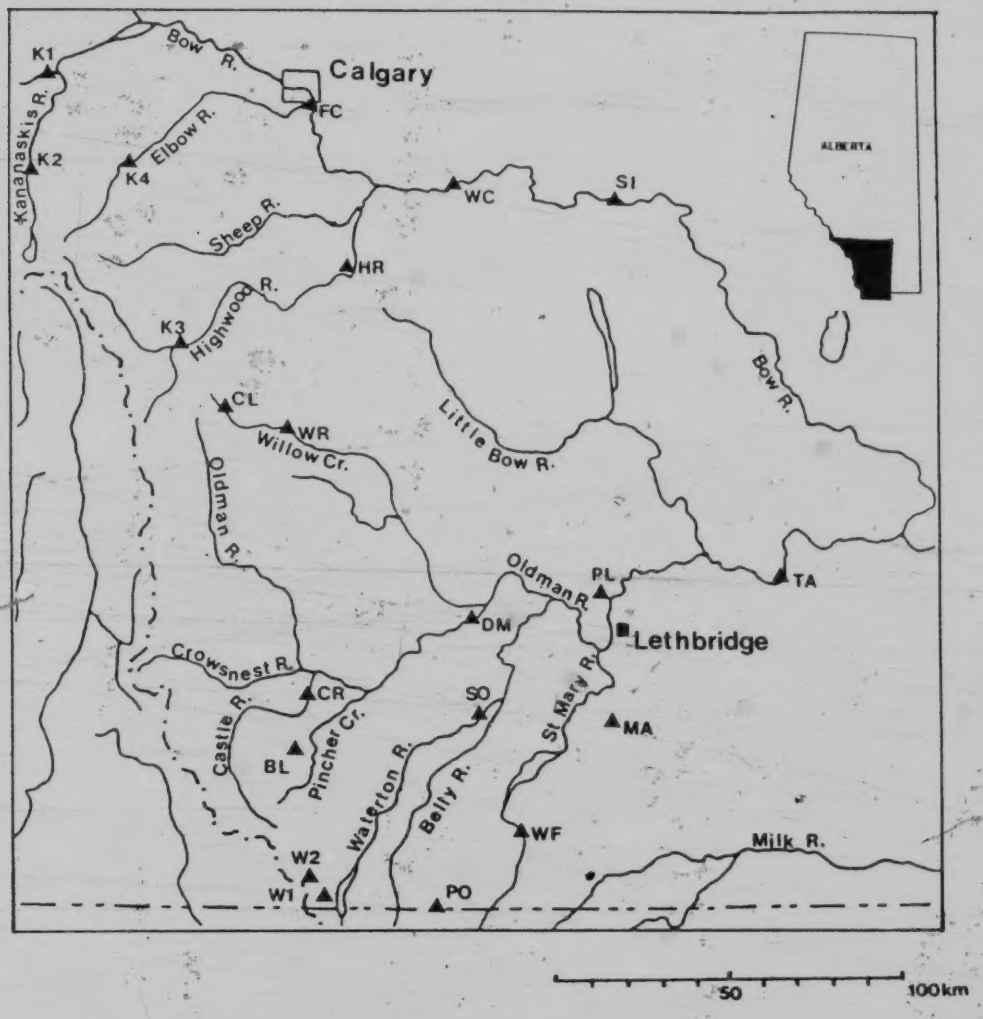


Figure 7 - Study area and 21 selected *P. angustifolia* and *P. balsamifera* locations (▲) in southwestern Alberta.

each location, 20 trees or fewer were selected for a total of 366 trees. The trees were selected to best represent the topography of the location and also the morphological variation found in the location. Selected trees were at least 15 meters apart from each other to try to avoid sampling the same multiple-stem clone. Each tree was tagged to allow further sampling. These 366 trees comprise the general sample which will be studied using morphology, flavonoid chemistry and electrophoresis.

3 - List of selected locations

The letter in parenthesis after each location name is that used in figures 13, 20 and 24.

1 - Beauvais Lake - BL (B)
Beauvais Lake Provincial Park, on highway 507, 25km South West of Pincher Creek., Beaver Creek day use area, near Beaver Creek and Beauvais Lake. Balsam poplars and aspens.
20 trees sampled.

2 - Chain Lakes - CL (C)
Private farm, 2km West of highway 22 and 2 km South of Chain Lakes Provincial Park entrance. Along Willow Creek. Balsam poplars and aspens.
20 trees sampled.

3 - Castle River - CR (R)

Castle River Recreation Area, 6km West of Pincher Creek, South of highway 3. Along Castle River. Balsam poplars and aspens.
20 trees sampled.

4 - Daisy May - DM (D)
Daisy May Campground, 249 Lyndon Road in Fort Macleod, North of highway 3. Riverside camping area along Old Man River. Balsam poplars and willows.
19 trees sampled.

5 - Fish Creek - FC (F)
Fish Creek Provincial Park in South Calgary. Hull's Wood Picnic Area near Bow River and Fish Creek, access via Bow Bottom Trail. Balsam poplars only.
20 trees sampled.

6 - High River - HR (H)
George Lane Memorial Park in High River on Macleod Trail. In campground along Highwood River. Balsam poplars only.
20 trees sampled.

7 - Kananaskis 1 - K1 (I)
Bow Valley Provincial Park, 12km East of Canmore on Trans Canada highway. In campground area, along Bow River. Balsam poplars, aspens and conifers.
20 trees sampled.

8 - Kananaskis 2 - K2 (J)

Kananaskis Country, on highway 40, 30 km South intersection with Trans Canada highway. Ribbon Creek Recreation Area, in picnic area along Ribbon Creek (Kananaskis River tributary). Balsam poplars, aspens and conifers.
20 trees sampled.

9 - Kananaskis 3 - K3 (K)
Kananaskis Country, on highway 541, 40km West of Longview. Sentinel Recreation Area, along Cataract Creek (Highwood River tributary). Balsam poplars, aspens and conifers.
20 trees sampled.

10 - Kananaskis 4 - K4 (L)
Kananaskis Country, on highway 66, 15km West of Bragg Creek. Allen Bill Pond Recreation Area, day use area along Elbow River. Balsam poplars, aspens and conifers.
20 trees sampled.

11 - Magrath - MA (M)
Magrath Campground in Magrath on highway 62, near Pothole Creek. Balsam poplars and various planted species.
5 trees sampled.

12 - Park Lake - PL (P)
Park Lake Provincial Park, on highway 25, 20km North of Lethbridge. Small bush of *Populus angustifolia* on border of the Park, along highway, near the lake. Narrowleaf cottonwoods, plains cottonwoods and willows.
10 trees sampled.

13 - Police Outpost - PO (O)

Police Outpost Provincial Park, 20km South of Cardston, access on highway 2. In camping area, near Outpost Lake. Balsam poplars and aspens.

18 trees sampled.

14 - Siksika - SI (S)

Siksika Vacation Resort, on highway 842, 10km South of Cluny. Along Bow River. Balsam poplars only.

20 trees sampled.

15 - Stand Off - SO (U)

Stand Off Recreation Area on highway 2, 30km South of Fort Macleod. Along Waterton River. Narrowleaf cottonwoods only.

20 trees sampled.

16 - Taber - TA (T)

Taber Provincial Park, on highway 864, 5km North of Taber and junction with highway 3. In Day use area and group use area, along Oldman River. Mostly plains cottonwoods and some balsam poplars.

4 trees sampled.

17 - Waterton 1 - W1 (A)

Waterton National Park, Service Road to Coppermine Creek, 2km from Red Rock Canyon, near Crandel Campground. Along Coppermine Creek (Blakiston Creek tributary). Balsam poplars, aspens and conifers.

20 trees sampled.

18 - Waterton 2 - W2 (W)

Waterton National Park, on Akamina Parkway to Cameron Lake, 5km West of town site. Along Cameron Creek. Balsam poplars, aspens and conifers.

10 trees sampled.

19 - Wyndham-Carseland - WC (X)

Wyndham-Carseland Provincial Park, on highway 24, 10km South of Carseland. In campground area, along Bow River. Balsam poplars only.

20 trees sampled.

20 - Woolford - WF (Y)

Woolford Provincial Park, on highway 820, East of Cardston, South of highway 5. In playground area and picnic area, along St. Mary River. Balsam poplars only.

20 trees sampled.

21 - Willow Creek - WR (Z)

Willow Creek Provincial Park, on highway 527, 20km South of Stavely and highway 2. In campground area, along Willow Creek. Balsam poplars only.

20 trees sampled.

Table 5 indicates the altitude, latitude and longitude of the 21 locations.

Table 5 - Altitude, latitude, longitude of the 21 *Populus* locations selected for the study of balsam poplars in Southwestern Alberta.

Location	Altitude	Latitude	Longitude
BL(B)	1350m	49°25′	114°05′
CL(C)	1350m	50°13′	114°12′
CR(R)	1200m	49°32′	114°02′
DM(D)	.930m	49°43′	113°24′
FC(F)	1000m	50°55′	114°01′
HR(H)	1020m	50°35′	113°50′
K1(I)	1290m	51°05′	115°05′
K2(J)	1440m	50°56′	115°09′
K3(K)	1500m	50°25′	114°35′
K4(L)	1500m	50°53′	114°42′
MA(M)	.960m	49°24′	112°57′
PL(P)	.900m	49°48′	112°55′
PO(O)	1350m	49°00′	113°27′
SI(S)	.810m	50°46′	112°51′
SO(U)	.960m	49°30′	113°20′
TA(T)	.750m	49°49′	112°10′
W1(A)	1500m	49°06′	113°58′
W2(W)	1500m	49°04′	113°58′
WC(X)	.900m	50°50′	113°25′
WF(Y)	1100m	49°11′	113°11′
WR(Z)	1020m	50°05′	113°45′

The letter between brackets after the location initials is the code used on correspondence analysis plots in figures 13, 20 and 24.

4 - Additional locations

To properly study the putative hybrid individuals, it is necessary to compare them with the pure species involved. For each taxon, collection sites situated outside the putative hybridization area were selected and 5 to 20 trees were sampled in each of them :

1. - *P. angustifolia* was collected at Gunnison, Colorado for morphology and chemotaxonomy
2. - *P. balsamifera* subsp. *balsamifera* was collected at Plamondon and High Level, Alberta for morphology and chemotaxonomy; at Plamondon, Lawrence Lake and Smoky River, Alberta for pollen viability and allozyme electrophoresis
3. - *P. balsamifera* subsp. *trichocarpa* was collected at Savona and Boston Bar, British Columbia for morphology and chemotaxonomy; at Burns Lake, Francois Lake and Vanderhoof, British Columbia for pollen viability and allozyme electrophoresis

Additional collections of *P. balsamifera* were also made in intermediate locations between the putative hybrid swarms of southwestern Alberta and the pure subspecies (figure 8) :

1-Cold Lake - CO

Cold Lake Provincial Park, 5km North East of Grand Centre on highway 55.

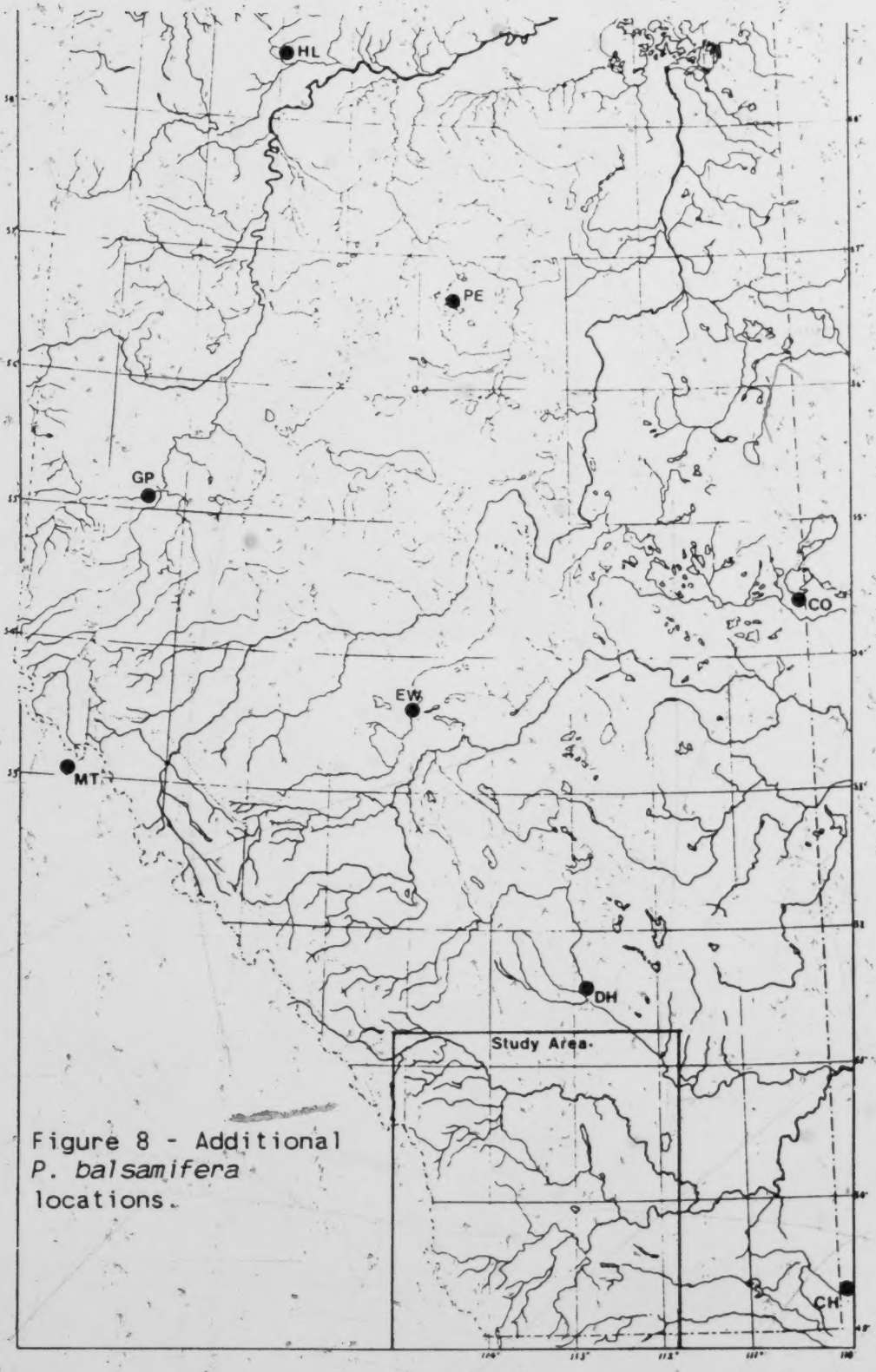


Figure 8 - Additional *P. balsamifera* locations.

2-Cypress Hills (Saskatchewan) - CH

Cypress Hills Provincial Park, 30km South of Maple Creek on highway 21.

3-Drumheller - DH

Midland Provincial Park, 5km West of Drumheller on highway 575.

4-Entwistle - EW

Pembina Provincial Park, near Entwistle, on Pembina River.

5-Grande Prairie - GP

5km South of Grande Prairie, on highway 40, on Wapiti River.

6-High Level - HL

14km East of High Level, on highway 58.

7-Mount Robson (British Columbia) - MT

20km West of Tête Jaune Cache on highway 16, at entrance of Mount Robson Provincial Park.

8-Peerless Lake - PE

8km North of Peerless Lake settlement.

5 - Methods : correspondence analysis

a - Method presentation

The three methods used produce large size multivariate sets of data. The type of data is different for each method :

1. - Morphological data are quantitative, the variables representing leaf measurements of an

individual.

2. - In flavonoid chemistry, the variables are binary, representing the presence or absence of a specific compound in an individual.

3. - Electrophoresis produces categorical variables, a numerical value being assigned to the alleles present at a certain locus for an individual.

With such large sets of data, multivariate analysis is helpful. It allows one to explore simultaneously all variables and individuals, to detect the underlying structure of the data and to reveal its characteristic features and patterns.

The multivariate analysis chosen in this study is correspondence analysis. This technique was first introduced by Hirshfield (1935) but ignored for many years. It was recently rediscovered and popularized by several authors (Benzecri 1969; Hill 1973, 1974; Tiel 1975). A complete description of the method is given by Greenacre (1984) and Lebart et al. (1984).

Correspondence analysis is analogous to PCA (principal component analysis), but appropriate to categorical data rather than continuous data (Hill 1982). First applied to contingency tables, it can however also be used with any matrix of categorical data, as well as continuous data, including individual by attribute matrices. Variables measured on a continuous scale can be transformed to a

discrete scale and treated as categorical data (Hill 1974, Greenacre 1984, Hoffman and Franke 1986).
As in PCA, correspondence analysis condenses most of the variation contained in directly observed variables, into a small number of uncorrelated variables called principal axes that are easier to interpret. Both methods are used for data display. The scores of each individual on the components or axes (generally the first two or three) can be calculated and plotted as a scatter diagram.

The unique feature of correspondence analysis is that one can view simultaneously both the rows (individuals) and the columns (characters) of the data matrix in the same low-dimensional space defined by the principal axes. This is possible because of the correspondence between the row components and the column components. Therefore, a simultaneous display of both the row variation and the column variation in the data matrix is produced and studied. The differences between PCA and correspondence analysis reside in the type of transformation applied to the original data matrix. In correspondence analysis, the original data matrix is rescaled e.g. all row and column sums are equal to 1. In addition the metrics used to calculate the distances between points (individuals or characters) in correspondence analysis are different from the ones used in PCA (Tiel 1975). Because of these mathematical differences, the terminologies of the two methods are different (Greenacre and Degos 1977) :

PCA	correspondence analysis
total variance	total inertia
covariance matrix	inertia matrix
principal components (independant variables)	principal axes
principal component scores	principal coordinates

Because of the rescaling of the data matrix, correspondence analysis illustrates variability in shape only, while PCA illustrates variability in both size and shape (Greenacre 1984).

b - Correspondence analysis example

The computer program (appendix 1) used in this study was written by Wayne Watson, statistical consultant at the University of Alberta. This program uses S.A.S. procedures (1985) and is based on the model described by Hoffman and Franke (1986).

To illustrate the interpretation of correspondence analysis, a small data matrix (table 6) was chosen as an example. This data matrix represents the presence or absence of 8 chemicals in the leaf extracts of 10 poplar trees. Using classical morphological characters, these 10 trees were identified as taxon 1, taxon 2 or putative hybrid (table 6).

Table 6 - Correspondence analysis example : presence-absence table of 8 chemicals (02 to 33) in leaf extracts of 10 poplar trees (T10 to T65).

---	02	06	08	12	13	16	17	33	- identification by morphology
T10	1.	1.	1.	0.	0.	1.	1.	1.	- taxon 1
T18	1.	1.	1.	1.	0.	0.	1.	1.	- taxon 1
T19	1.	1.	1.	0.	0.	0.	1.	1.	- taxon 1
T20	0.	1.	1.	1.	1.	0.	0.	0.	- taxon 2
T22	1.	1.	0.	1.	0.	0.	1.	0.	- taxon 1
T30	1.	1.	1.	1.	1.	1.	1.	1.	- putative hybrid
T36	1.	1.	1.	1.	0.	1.	1.	1.	- taxon1
T37	1.	1.	0.	1.	1.	0.	1.	0.	- taxon 2
T40	0.	1.	1.	1.	1.	0.	1.	0.	- taxon 2
T65	1.	0.	1.	0.	0.	1.	1.	1.	- taxon 1

In correspondence analysis, the overall variation in each set of points is called total inertia. It is calculated from the distances between rows or columns; the inertia being equivalent for both sets of points.

The total inertia is decomposed in terms of eigenvalues. Eigenvalues are the basic values (characteristic or latent roots) of the inertia matrix calculated after transformation of the original data matrix. They indicate the inertia explained by the principal axes.

In the example, the first eigenvalue is equal to 0.23 (table 7a) which represents 54.63% of the total eigenvalue sum. Therefore, the first principal axis accounts for 54.63% of the variation in the data. The second eigenvalue is equal to 0.09, the second principal axis accounting for 21.13% of the variation, etc.

Scores of the individuals (trees) and of the characters (chemicals) on the principal axes are calculated (table 7b and 7c). As the inertia is equivalent for both sets of points, the two sets of points can be represented simultaneously on the same graph. Thus, in figure 9, both trees and chemicals are plotted.

Distances between points of each set can be compared. For example, tree #18 and tree #19 are close on the graph which means a similar chemical pattern. It is also possible to interpret the position of one point of one set relative to all the points of the other set. However, the distances between individual points of different sets (trees and

Table 7 - Correspondence analysis example : presence-absence of 8 chemicals in 10 poplar trees.
a - Inertias (eigenvalues), percentages of total inertia and cumulative percentages
b - Principal coordinates for rows (trees)
c - Principal coordinates for columns (chemicals)
(for the first 5 principal axes)

AXIS#	EIGENVA	PERCT	CUMPCT
AXIS1	0.22969	54.63	54.63
AXIS2	0.08886	21.13	75.76
AXIS3	0.04100	09.75	85.51
AXIS4	0.03365	08.00	93.52
AXIS5	0.01398	03.32	96.84

a

ID	AXIS1	AXIS2	AXIS3	AXIS4	AXIS5
T10	+0.52846	+0.15594	-0.03966	+0.03987	-0.06429
T18	+0.05302	-0.30041	-0.20434	+0.02864	+0.15865
T19	+0.30550	-0.23510	-0.41951	-0.21149	-0.05290
T20	-0.98734	+0.37885	-0.17812	+0.11021	+0.15292
T22	-0.17841	-0.72323	+0.22060	+0.29734	-0.12031
T30	-0.04044	+0.24591	+0.18847	-0.04508	+0.01889
T36	+0.28020	+0.04410	+0.09051	+0.20979	+0.11867
T37	-0.34149	-0.20236	+0.25666	-0.38477	+0.03535
T40	-0.73005	+0.16715	-0.12364	+0.07343	-0.24516
T65	+0.74160	+0.30445	+0.06568	-0.03004	-0.08005

b

ID	AXIS1	AXIS2	AXIS3	AXIS4	AXIS5
C02	+0.3517	-0.29801	+0.09780	-0.06525	+0.01481
C06	-0.2575	-0.17487	-0.11469	+0.07144	+0.00172
C08	+0.0394	+0.31906	-0.38309	+0.11947	+0.00713
C12	-0.5796	-0.18689	+0.17648	+0.22550	+0.14381
C13	-1.0951	+0.49443	+0.17702	-0.33554	-0.08034
C16	+0.7876	+0.62932	+0.37657	+0.23786	-0.01433
C17	+0.1434	-0.20260	+0.01909	-0.01352	-0.21724
C33	+0.4551	+0.00601	-0.04386	-0.30612	+0.16231

c

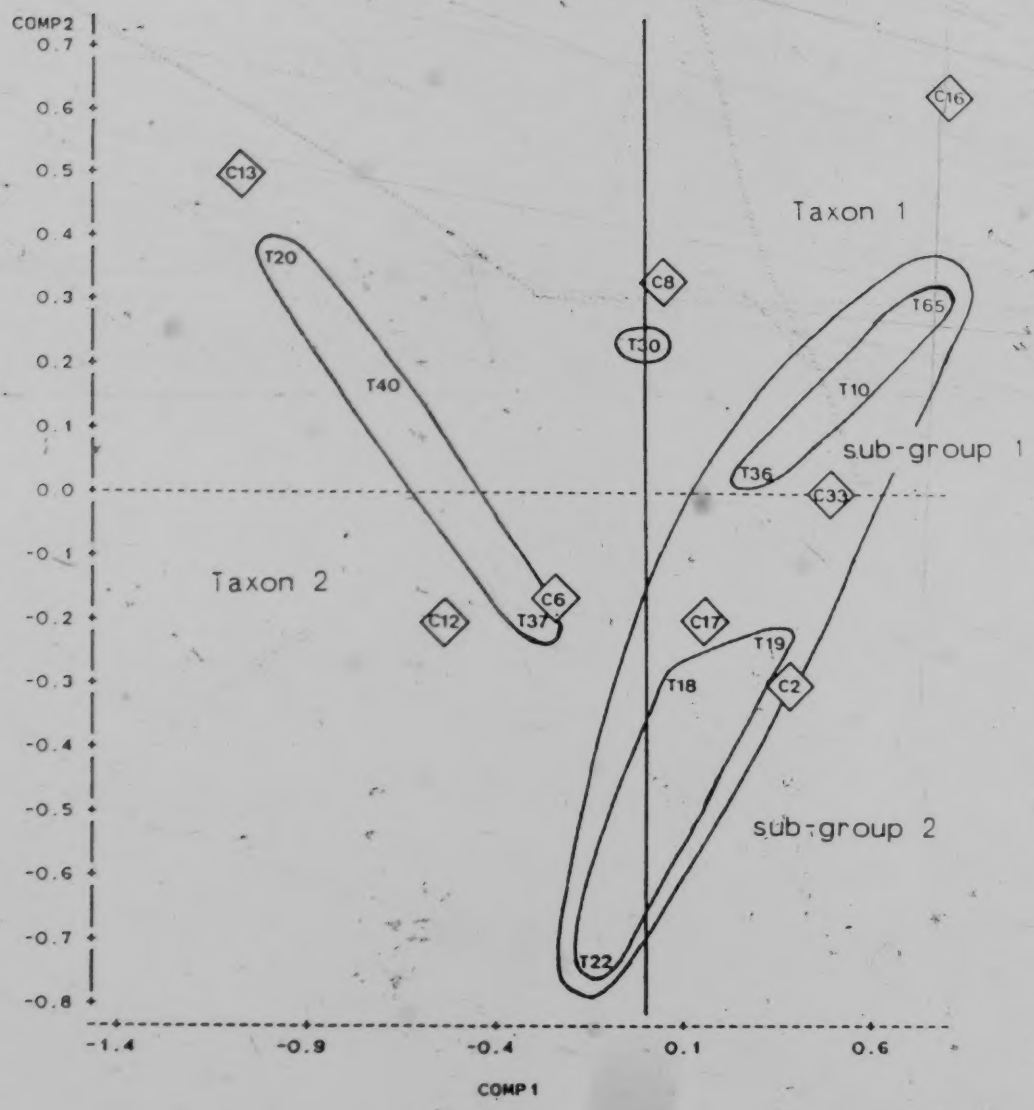


Figure 9 - Correspondence analysis example : presence-absence of 8 chemicals in 10 poplar trees, first 2 principal axes plot.

chemicals : C2 to C33 (in ◇)
trees : T10 to T65

chemicals) cannot be compared.
The center of gravity, located at the origin of the axes, corresponds to the average pattern of both sets of points. Points situated near the center such as chemical C6, have a less differentiated distribution among the trees than points situated far from the center such as chemical C13 and C16. Therefore, chemicals C13 and C16 are very useful as discriminators among the 10 trees of the sample. Chemical C6 is less useful.

Correspondence analysis also define two types of coefficients which determine if the results give a satisfactory representation of the information contained in the data (Lebart et al. 1984) :

1 - The importance of each point on the principal axis direction is quantified by their absolute contributions (table 8a). For each principal axis absolute contributions represent the percentage of inertia explained by each point. On the first principal axis, accounting for 54.63% of the total inertia, 18% of this inertia is explained by chemical C12 (AC=0.18). As well, chemicals C13 (AC=0.37), C16 (AC=0.19) and C33 (AC=0.11) also have high absolute contributions. These four chemicals contribute to the direction or nature of axis 1. Chemicals C13 (AC=0.19) and C16 (AC=0.32) contribute to axis 2.
As the total inertia is equivalent for both sets of points,

Table 8 - Correspondence analysis example : presence-absence of 8 chemicals in 10 poplar trees.
a - Absolute contributions of row and column points to inertia for the first 5 principal axes (contributions sum to 1 within a principal axis)

ID	AXIS1	AXIS2	AXIS3	AXIS4	AXIS5	rows(trees)
T10	0.130269	0.029319	0.004109	0.005061	0.031681	
T18	0.001311	0.108810	0.109101	0.002612	0.192942	
T19	0.036278	0.055535	0.383212	0.118676	0.017875	
T20	0.303149	0.115368	0.055264	0.025784	0.119507	
T22	0.009899	0.420441	0.084772	0.187664	0.073973	
T30	0.001017	0.097218	0.123756	0.008628	0.003648	
T36	0.042726	0.002735	0.024975	0.163484	0.125936	
T37	0.054397	0.049372	0.172131	0.471387	0.009578	
T40	0.207173	0.028073	0.033286	0.014308	0.383924	
T65	0.213781	0.093130	0.009394	0.002395	0.040935	

ID	AXIS1	AXIS2	AXIS3	AXIS4	AXIS5	col(chemicals)
C02	0.076926	0.142776	0.033321	0.018073	0.002243	
C06	0.046383	0.055303	0.051551	0.024374	0.000034	
C08	0.000964	0.163655	0.511287	0.060599	0.000519	
C12	0.182826	0.049133	0.094947	0.188891	0.184953	
C13	0.372914	0.196502	0.054586	0.238987	0.032983	
C16	0.192886	0.318345	0.247015	0.120094	0.001050	
C17	0.014381	0.074235	0.001428	0.000873	0.542622	
C33	0.112720	0.000051	0.005865	0.348110	0.235597	

chemicals and trees, it is also possible to describe the principal axes with the absolute contributions of the trees. When considering the trees, axis 1 is mainly defined by trees #10, 20, 40 and 65. Trees #18, 20 and 22 contribute to axis 2.

2 - The quality of the representation on each point in the graph is determined by the relative contributions (table 8b). A high relative contribution of an axis to a point indicates that this axis explains the point very well. Axis 1 contributes highly to the inertia of trees #10, 19, 20, 36, 37, 40 and 65. Axis 2 explains the inertia of trees #18, 22 and 30. All points of both sets are well explained by the two first principal axes, which cumulate 75.76% of the total inertia.

The first principal axis separates trees which exhibit chemicals 16 and 33 from trees exhibiting chemicals 12 and 13 (figure 9). The second principal axis separates trees with chemicals 13 or 16 from trees without any of these chemicals.

The first dimension separates taxon 1 from taxon 2. Taxon 2 is characterized by the presence of chemicals 12 and 13 and the absence of chemicals 16 and 33. Taxon 1 is characterized by the absence of chemical 13. The second dimension distinguishes two sub-groups in taxon 1 (figure 9). The first sub-group, including trees # 10, 36 and 65, is

Table 8 - Continued.
b - Relative contributions of row and column points to inertia for the first 5 principal axes (contributions sum to 1 across principal axes)

ID	AXIS1	AXIS2	AXIS3	AXIS4	AXIS5	rows(trees)
T10	0.807894	0.070344	0.004549	0.004598	0.011956	
T18	0.015701	0.504124	0.233247	0.004582	0.140606	
T19	0.247044	0.146307	0.465860	0.118394	0.007407	
T20	0.811430	0.119466	0.026407	0.010111	0.019465	
T22	0.044502	0.731264	0.068036	0.123600	0.020237	
T30	0.016241	0.600561	0.352771	0.020184	0.003545	
T36	0.516767	0.012799	0.053926	0.289677	0.092689	
T37	0.312264	0.109646	0.176396	0.396423	0.003346	
T40	0.810256	0.042475	0.023240	0.008198	0.091371	
T65	0.798340	0.134547	0.006263	0.001310	0.009302	

ID	AXIS1	AXIS2	AXIS3	AXIS4	AXIS5	col(chemicals)
C02	0.514023	0.369088	0.039747	0.017692	0.000912	
C06	0.433026	0.199744	0.085917	0.033336	0.000019	
C08	0.005745	0.377324	0.543962	0.052908	0.000188	
C12	0.689736	0.071710	0.063945	0.104398	0.042460	
C13	0.751450	0.153188	0.019636	0.070551	0.004044	
C16	0.508069	0.324404	0.116153	0.046342	0.000168	
C17	0.173415	0.346324	0.003075	0.001542	0.398181	
C33	0.618860	0.000108	0.005748	0.279991	0.078711	

characterized by the absence of chemical 16. The second sub-group, including trees #18, 19 and 22, is characterized by the presence of this chemical. The hybrid status of tree # 30 is confirmed by its intermediate position between taxon 1 and taxon 2. In this individual, chemicals 12, 13, 16 and 33 are present.

Also, the position of tree #37 is unclear. Identified as Taxon 2 with morphological characters, this tree seems to have an intermediate position between Taxon 1 and 2 on the graph. This position is explained by the absence of chemical 8 in this tree which is also absent from tree #22 belonging to Taxon 1. Chemical 8 contributes to the direction of axis 3 (AC = 0.51). Tree #37 should not be interpreted as a putative individual. Contrary to tree #30, it presents the typical chemical pattern of Taxon 2. However, chemical 8 should be interpreted as a "fluctuating" chemical without taxonomical significance.

c - Data structure considerations

As stated earlier, the three methods used in this study produce different types of data :

- Morphological data are continuous; to fit the correspondence analysis requirements, they were transformed into discrete data (appendix 2);
- Flavonoid chemistry data are presence-absence data (appendix 3);
- Electrophoresis data are categorical (appendix 4).

Correspondence analysis is first employed with contingency tables and binary data (Lebart et al. 1984). It was then necessary to check if the three data matrices are suited to correspondence analysis. Correspondence analysis was therefore performed on the three data sets, before and after two types of recoding :

- Before recoding, the data sets (appendices 2, 3 and 4) were directly analysed without modifications : original data plots;
- After recoding,

  1 - the data sets were recoded by doubling in dummy variables to create a "pseudo" contingency table (Greenacre 1984, ch.6) : doubled data plots;

  2 - the data sets were recoded using binary coding (Lebart and al. 1984, ch.4) : binary data plots.

When compared, the different types of plots appeared very similar. For a same data set, the original data plot and the doubled data plot presented slight differences, only in principal axes scaling. The binary data plot was a mirror image (first axis direction inverted) of the former ones. On the three types of plots, the distribution (or display) of the individuals was rigorously the same. Therefore, for computational reasons as well as clarity of plot interpretation, the original data sets were chosen to perform correspondence analysis, as illustrated in the example presented in the precedent paragraph.

Also, in the case of the morphological data set, a principal component analysis was performed on the original continuous data. Correpondence analysis was performed on the same data but after transformation into discrete data. Again, correspondence analysis and PCA plots were very similar. For reasons of technique homogeneity, the three data sets were studied by correspondence analysis.

IV - Morphology

1 - Definition

"Plant morphology is the study of form and structure of plants and, by implication, an attempt to interpret these on the basis of similarity of form and origin." (Bold 1967)

2 - Literature review of the morphological studies on North American poplars

Most morphological studies are based on leaf and twig characters. Morphology was studied either alone, or in combination with other techniques.

Using morphological characters alone, Brayshaw (1965b) surveyed southern Alberta to delimit the range of *Populus* species and putative hybrid combinations present in the area. He also described these taxa in detail. Rood et al.(1986) studying the same area of southern Alberta found continuous variation in leaf shape among the 56 clones studied. They concluded that the three species *P. angustifolia*, *P. balsamifera* and *P. deltoides* form a single interbreeding population in southern Alberta. Barnes (1969, 1975) investigated intra- and interclonal variation in *P. tremuloides* and *P. grandidentata* and found a significant

amount of interclonal variation. This should allow the distinction of different clones, but the intraclonal variation must be clearly understood and controlled for the technique to be efficient.

Crawford (1974) studied morphological and flavonoid characters to assess the hybrid status of *P. acuminata*.

Eckenwalder (1977) used morphology, chromosome numbers and flavonoids to clarify the systematics of *Populus* in southwestern North America.

Hu et al. (1985) found that vegetative characters were generally sufficient to establish a satisfactory classification of 13 *Populus* taxa when using character means. However, because of individual variation, it was not possible to assign an individual to a particular taxon.

Comtois et al. (1986) used morphological characters and isozymes to study clonal constitution and sex ratio in *P. balsamifera*. The morphological results were found unsatisfactory to distinguish clones because of high intraclonal variability.

Sokal et al. (1986) studied geographic variation of *P. deltoides* and suggested glaciation, small scale micro-climatic and site variation were possible factors responsible for the pattern observed, rather than global, continent-scale climatic factors.

3 - Materials and techniques

a - Selection of material

All the morphological studies on *Populus* emphasize the important intra-individual variation of morphological characters in the genus. Hu et al. (1985) attributed part of this variation to environmental and developmental plasticity. Also, as stated in chapter II-3, *Populus* exhibits leaf dimorphism (Critchfield 1960). "Early" and "late" leaves are produced, depending on the time of their initiation and differentiation. Consequently, the sampling method of the leaf material is critical when using leaf and stem characters for taxonomic purposes, especially at the subspecies and variety levels.

Barnes (1969) remarked that "late" leaves exhibit more intra-individual variation than "early" leaves. Figure 10 shows this important variation in a set of late leaves from a single twig of *P. balsamifera* (Vanende 1984). Variation in both shape and dimension is involved. Furthermore, leaf characters show more intraspecific variation in the early leaves than in the late leaves (Eckenwalder 1977). Therefore, only early leaves were selected for the study of morphological characters.

b - Collection and leaf selection

Short shoots, which bear only early leaves, were used for

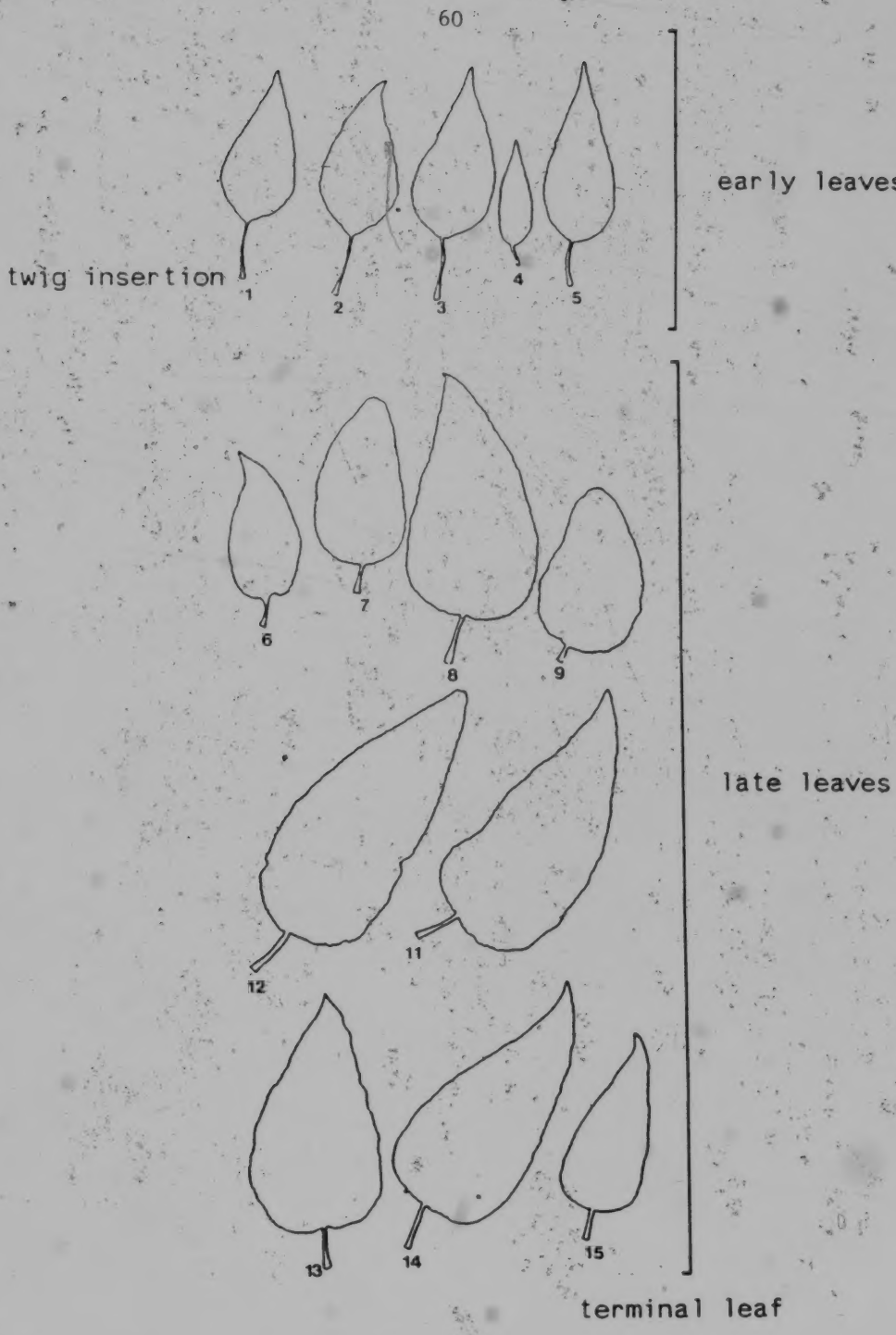


Figure 10 - Dimension and shape variability of the "late leaves" (#6 to 15) of *P. balsamifera* on a one-year-old twig (Vanende 1984).

all collections of the general sample.
The collection was made in July 1986. To determine the time of the year when the leaves reach their maximum size, short shoots were collected on a tree located in Edmonton (Saskatchewan River Valley), every week, from May 12th to October 29th 1987. The leaves reached their maximum size in mid June. It was assumed that by the end of July, the collecting period for the general sample, leaves of the poplar populations of southwestern Alberta had also reached their maximum size. However, many of the collection sites suffered serious late frosts during the spring. Many buds were killed and the production of the two types of foliage was perturbed. Those which survived produced only long shoots. In some areas, short shoots were not available for collecting, especially in the foothills region. Elsewhere, mostly around Calgary, an aphid infestation damaged the foliage.
The following growing season was better, allowing the completion of the collection in July 1987.

On each of the 366 trees of the general sample, two short shoots were collected in the median part of the crown, as recommended by Brayshaw (1965b) and Barnes (1969). These short shoots were collected on two- or three-year-old twigs borne by main branches of the trunk. Suckers and adventitious shoots from the basal part of the trunk were avoided. The shoots were pressed and dried. Then the leaves,

generally 5 to 7, were pulled from the shoots and taped on herbarium mounting paper in order of appearance on the shoot (figure 11).

To record the morphological characters, the first or basal leaf on the short shoot was avoided. When present, this leaf was either badly damaged or atrophied and deformed. The next three leaves were observed to exhibit the minimum amount of intra-individual variation and were measured. The last leaves on the short shoot become progressively longer and narrower (as shown on figure 11a) and were also avoided.

c - Choice of characters

The characters studied on poplar leaves are very similar among authors (table 9). The different measurements include blade and petiole lengths, blade width and several angles (figure 12).

Other quantitative characters include : teeth number per leaf side, maximal tooth depth or size, number of leaves per shoot, leaf surface area and perimeter and number of basal glands (0, 1 or 2).

Qualitative characters include : leaf and twig colour, petiole shape (rounded or flattened cross-section), pubescence and bud shape.

In this study preference was given to quantitative characters. Most of the qualitative characters are too

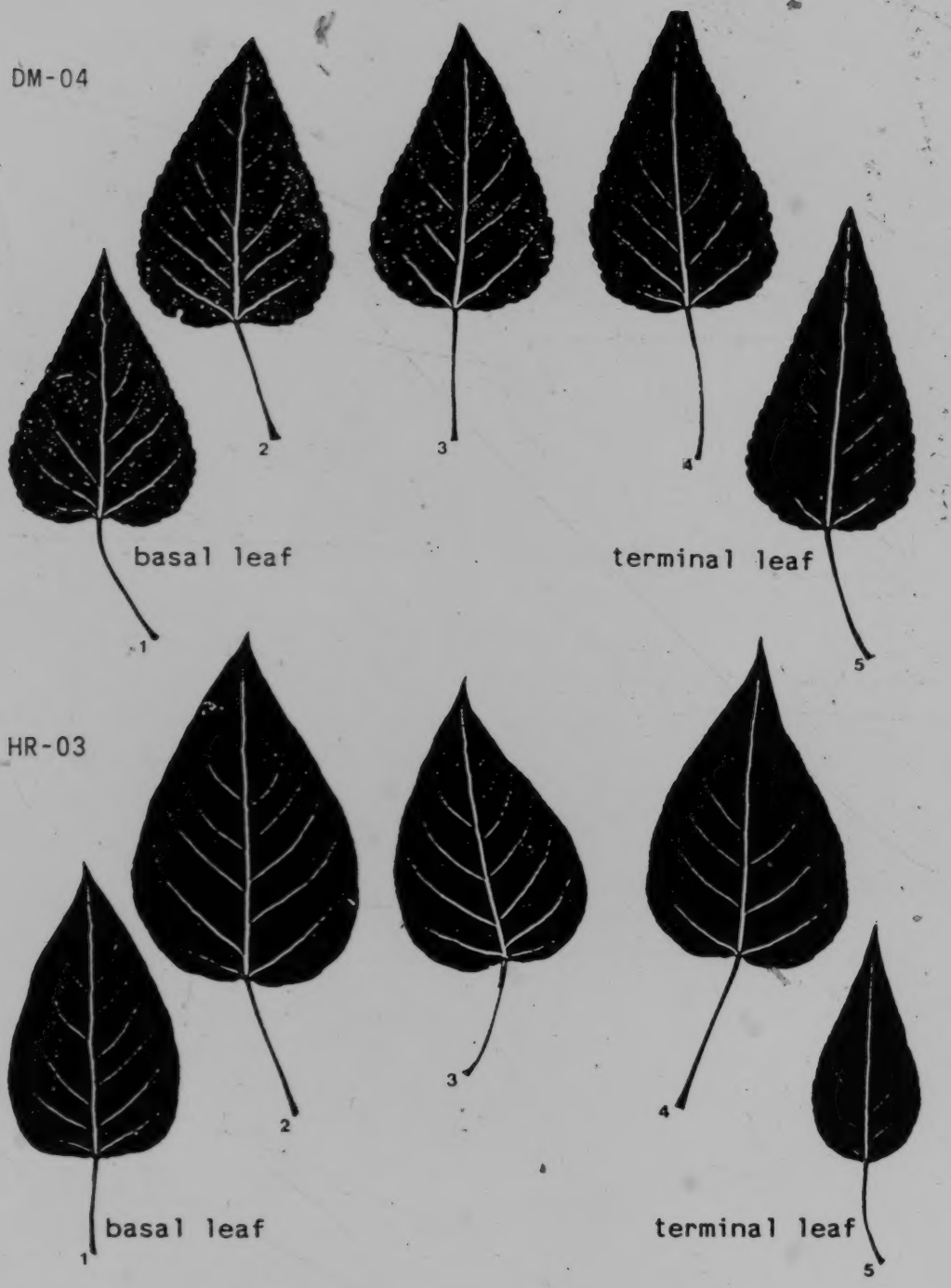


Figure 11a - Silhouettes of *Populus balsamifera* leaf samples from High River location (HR-03) and Daisy May location (DM-04) (the leaves are shown at 64% of their real size).

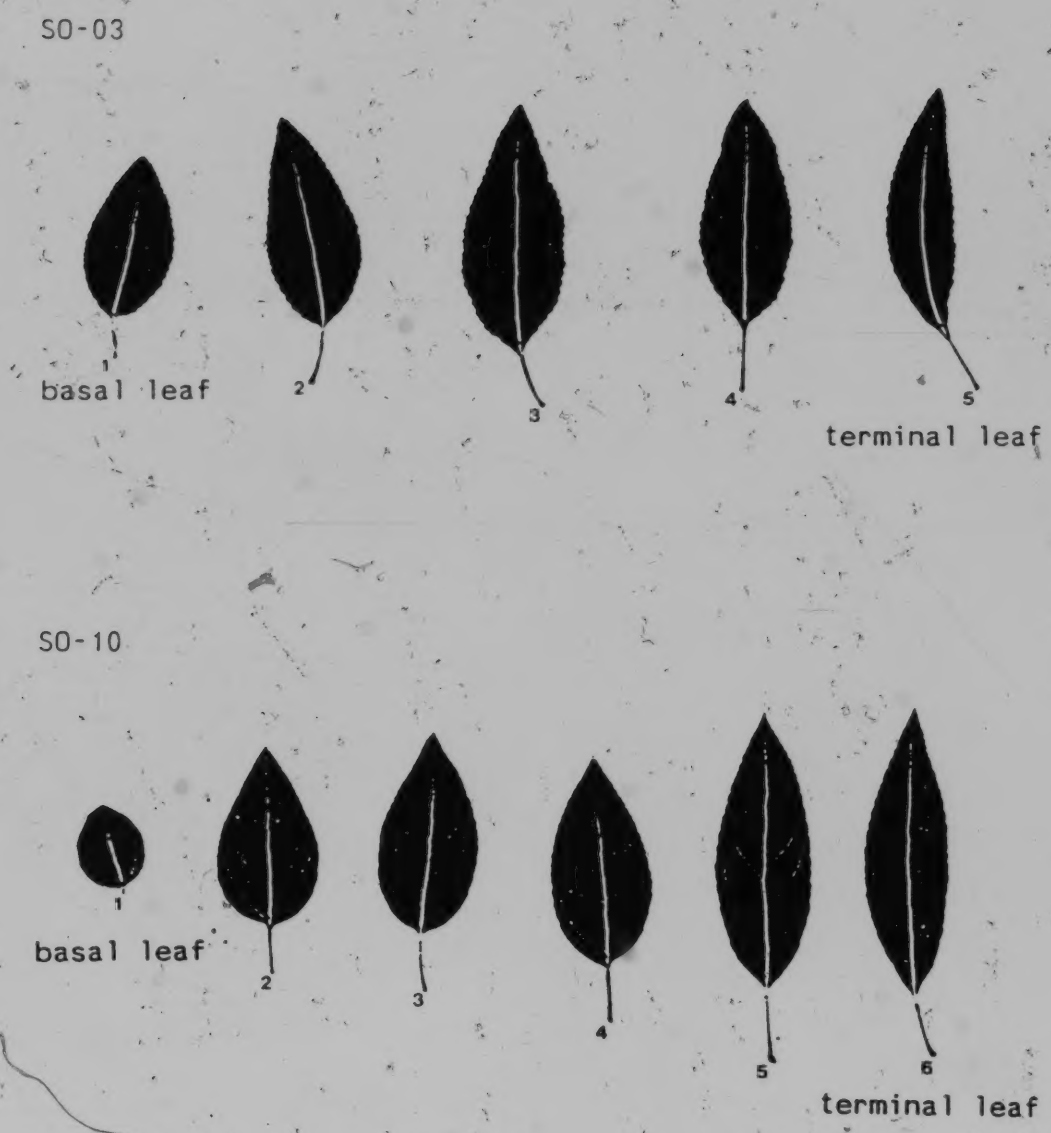


Figure 11b - Silhouettes of *Populus angustifolia* leaf samples from Stand Off location (SO) (the leaves are shown at 64% of their real size).

Table 9 - Morphological characters used in various studies of North American poplars.
L : leaf blade length; P : petiole length;
W : leaf blade width; D : distance to the maximum width;
T : number of teeth on one or both sides of leaf

Authors	-Leaf type -Species studied or area	Characters
Brayshaw (1965b)	-not specified -South western Alberta	Leaf apex, teeth and petiole shape, leaf and twig colour, P/L, W/L, D/L.
Barnes (1969)	-early leaves -*P. grandidentata P. tremuloides*	Pubescence, W, L, D, P, T, tooth size and irregularity, W/L.
Crawford (1974)	-not specified -*P. angustifolia, P.deltoides, P. x acuminata*	Petiole cross section, P, T, L/W, P/L.
Eckenwalder (1977)	-early leaves -South western North America	Leaf base shape, bud shape, leaf pubescence, tooth depth, L, W, P, T.
Hu et al. (1985)	-late leaves -13 taxa	Number of leaves on twig, twig length and width, leaf base and apex angles, internode, petiole shape, P, L, D, W.
Rood et al. (1985)	-late leaves (young trees) -South western Alberta	Leaf surface, tooth depth, apex angle, P, L, W, D, T.
Comtois et al. (1986)	-early leaves -*P. balsamifera*	Base and apex angles, leaf perimeter, P, L, W, D, T.
Sokal et al. (1986)	-late leaves -*P. deltoides*	Basal glands, twig length and diameter, internode, base and apex angles, P, W, D.

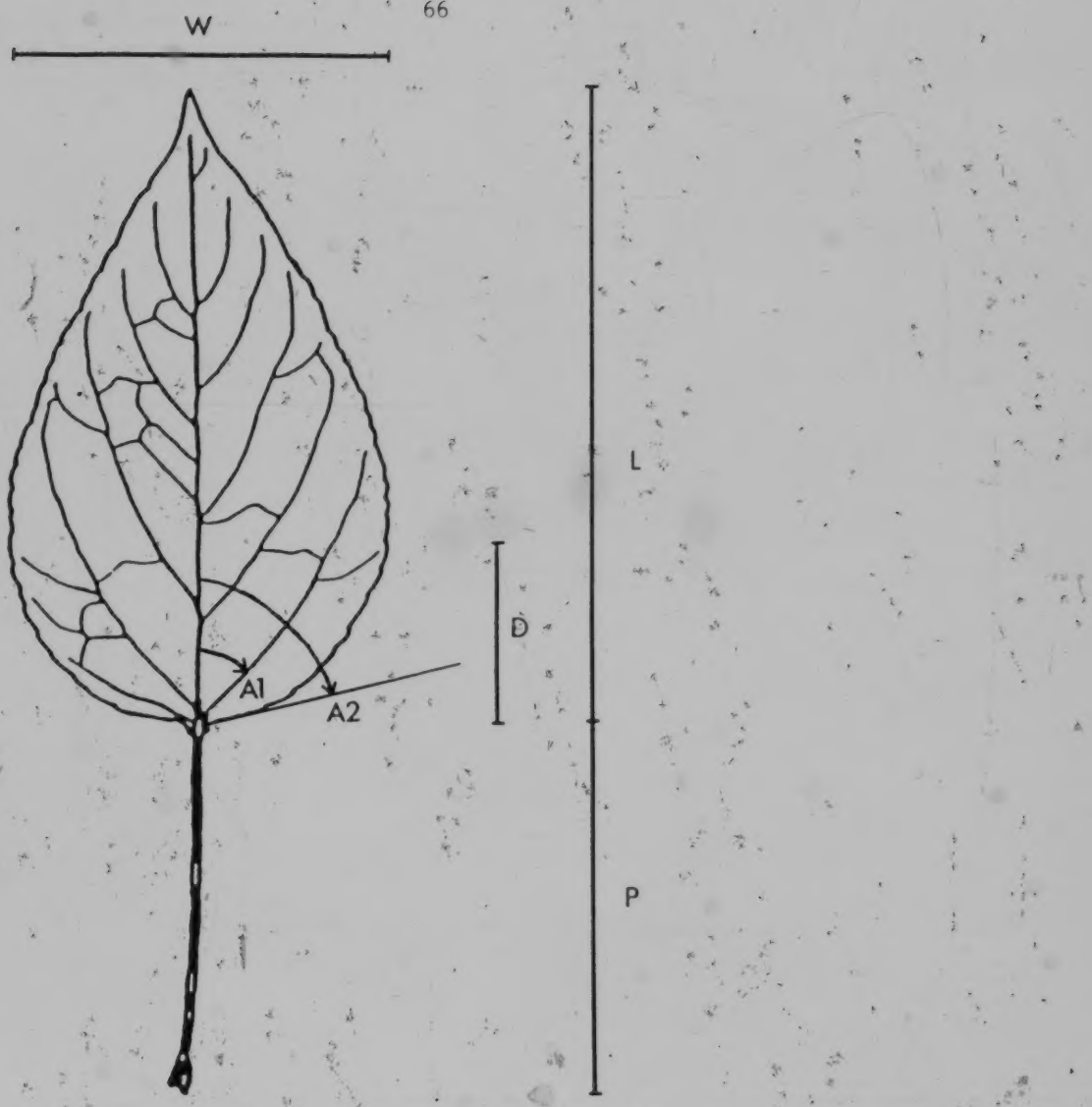


Figure 12 - Quantitative characters measured on the poplar leaves in this study.
- L : blade length
- W : maximum blade width
- P : petiole length
- D : minimum distance from blade insertion to point of maximum width
- T : total number of teeth on left side of blade
- A1: angle between main vein and basal or lower axillary vein
- A2: maximum angle between main vein and leaf base

subjective, especially when dealing with closely related taxa.

Seven quantitative characters were selected as follow (figure 12) :

- L : blade length
- W : maximum width of blade
- P : petiole length
- D : minimum distance from blade insertion to point of maximum width
- T : total number of teeth on left side of blade
- A1: angle between main vein and basal or lower axillary vein
- A2: maximum angle between main vein and leaf base

In the material collected, too many leaf apices were damaged and the apex angle character had to be discarded.

Also, in many leaves, one side was deformed by the presence of galls and it was not possible to record the presence of basal glands.

Tooth depth and size were not selected; those characters are useful mostly when considering *P. deltoides*, a taxon not part of this study.

Leaf surface area and perimeter were considered redundant with the other blade measurements and not utilized.

The seven characters were recorded for the six leaves (three leaves on two short shoots) selected for each tree of the general sample.

4 - Data treatment

The mean values of the six measurements for each character and for each tree were computed (appendix 2). In spite of the standardization of the collecting method, leaves of the same shape and characteristics may differ in size due to external factors such as climate. To avoid this bias, Brayshaw (1965), Barnes (1969) and Crawford (1974) used ratios (to the blade width or length) in combination with or instead of raw measurements, a procedure followed in the present study.
The original data matrix (appendix 2) was transformed by computation of ratios of each character to the blade length :

- W/L : maximum width/blade length
- P/L : petiole length/blade length
- D/L : distance to maximum width/blade length
- T/L : number of teeth/blade length

The two angle measurements were not modified. This transformation of the results reduced the number of characters to a total of six.

Before performing the correspondence analysis, the continuous data were transformed into discrete data by dividing the total range of each character into ten classes.

5 - Results

Correspondence analysis was performed on the morphological data matrix (figure 13).
The first axis accounts for 69% of the variation in the data. It can be defined as :
- the number of teeth (T/L, Absolute Contribution or AC = 0.32) combined with the distance to the maximum width (D/L, AC = 0.38)
versus
- the basal angle (A1, AC = 0.11) combined with the petiole length (P/L, AC = 0.08) and the blade width (W/L, AC = 0.07).
The second axis accounts for 15% of the variation in the data. It can be defined as the number of teeth (T/L, AC = 0.40) versus the distance to the maximum width (D/L, AC = 0.37).

During the collection of the material, the trees were tentatively identified using Brayshaw's leaf morphology criteria (1965a). Individuals were classified into three groups : those resembling *P. angustifolia*, those resembling *P. balsamifera* (the two subspecies could not be distinguished with these criteria), and intermediate individuals resembling putative hybrids between *P. angustifolia* and *P. balsamifera*.

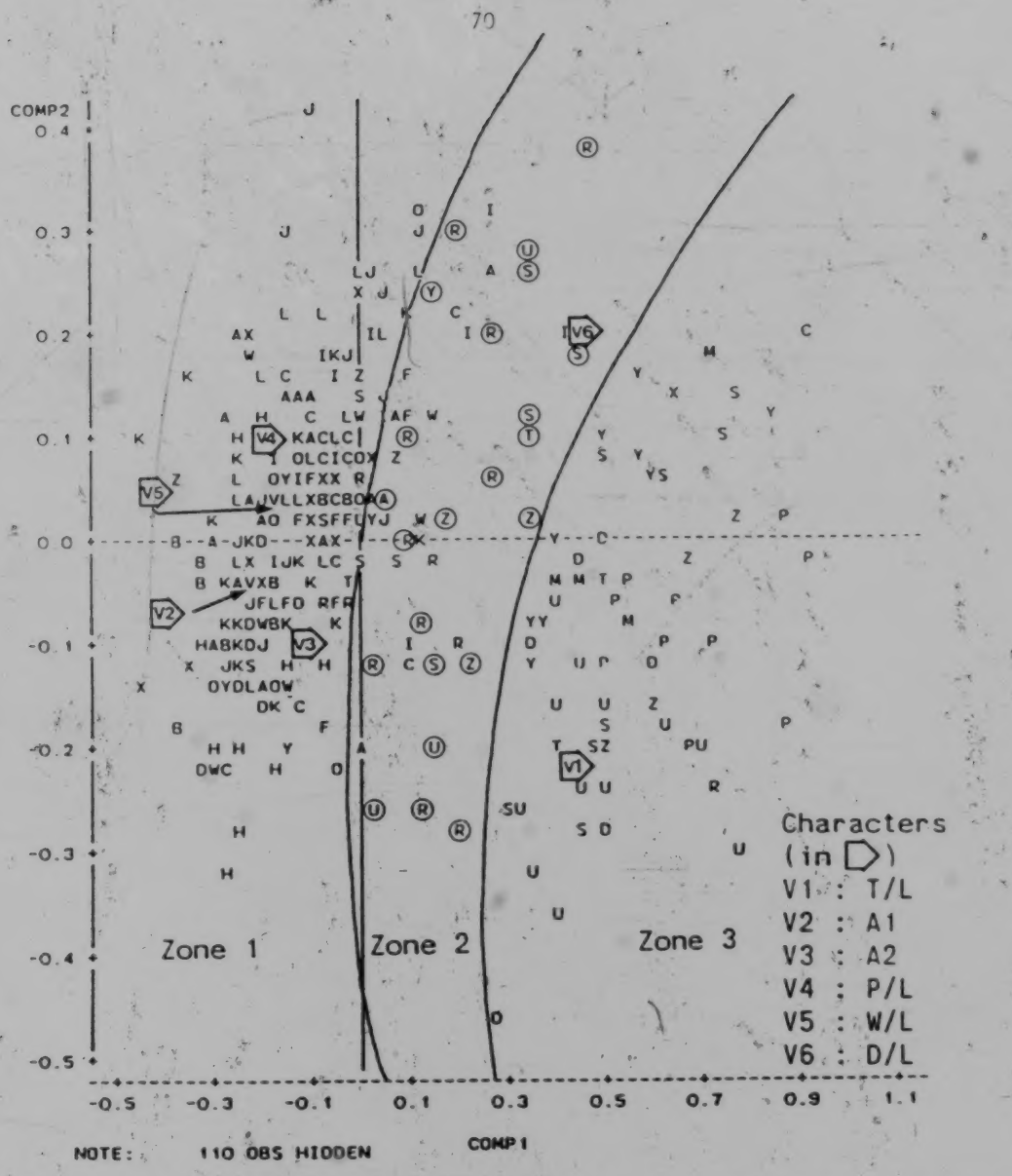


Figure 13 - Correspondence analysis plot (first 2 principal axes) of the transformed morphological data matrix.

- Each letter plotted corresponds to an individual and indicates its location code (location codes are listed in table 5 and chapter III-3).

- Zone 1 : all individuals identified as *P. balsamifera*
- Zone 2 : individuals identified as putative hybrids between *P. angustifolia* and *P. balsamifera* are circled. The non-circled individuals are all identified as *P. balsamifera*
- Zone 3 : all individuals identified as *P. angustifolia*

This identification was used to draw a zonation on the correspondence analysis plot. The zonation generally delimits the three groups on the plot, but there is some overlap. This lack of disjunction means that an almost continuous variation of leaf shape exists in the population studied. However, general trends can be distinguished (figure 14 and table 10).

A key can also be constructed (table 11):

- *P. angustifolia* individuals have high positive scores on axis 1, which means their leaves have : small basal angles, narrow blades and short petioles.
The number of teeth and distance to the maximum width variations are combined on axis 1 and 2, which means the leaves have : large number of teeth and long distances to the maximum width.

- *P. balsamifera* individuals have negative scores on axis 1, which means their leaves have : large basal angles, wide blades and long petioles.
The number of teeth and distance to the maximum width, combined on axis 1 and 2, permit the distinction of two extremes in *P. balsamifera* :

1. At one extreme are individuals which correspond to subspecies *balsamifera* (Brayshaw 1965b) : their leaves have long distances to maximum width and small numbers of teeth.
2. At the second extreme are individuals which

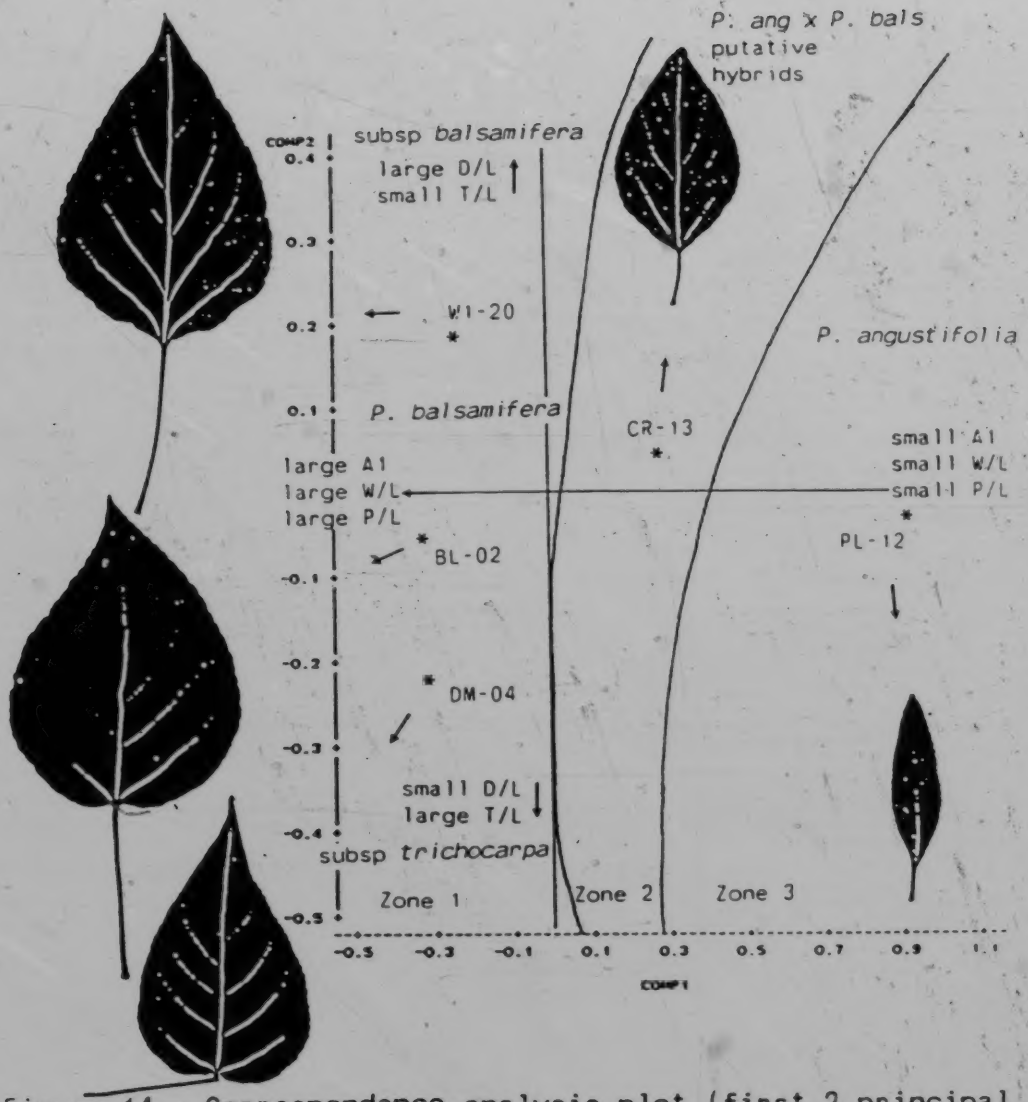


Figure 14 - Correspondence analysis plot (first 2 principal axes) of the transformed morphological data matrix;
The different trends of leaf shape exhibited in the population are illustrated with leaf examples.
- Zone 1, *P. balsamifera*
W1-20 from Waterton 1 location, at subsp. *balsamifera* extreme
DM-04 from Daisy May location, at subsp. *trichocarpa* extreme
BL-02 from Beauvais Lake location, intermediate
- Zone 2, *P. angustifolia* x *P. balsamifera* putative hybrid
CR-13 from Castle River location
- Zone 3, *P.angustifolia*
PL-12 from Park Lake location
(Character values for the five examples are in table 10)

Table 10 - Values of the morphological characters for the examples shown on figure 14; these values are the means calculated from the measurements done on six leaves per individual.

- PL-09 : *P. angustifolia*
- CR-13 : *P. angustifolia* x *P. balsamifera* putative hybrid
- W1-20 : *P. balsamifera* at subsp. *balsamifera* extreme
- BL-02 : *P. balsamifera* intermediate
- DM-04 : *P. balsamifera* at subsp. *trichocarpa* extreme

	PL-09	CR-13	W1-20	BL-02	DM-04
A1	40.8°	63.3°	68,3°	105.8°	110°
W/L	0.41	0.68	0.66	0.71	0.64
P/L	0.23	0.30	0.48	0.52	0.47
D/L	0.36	0.38	0.26	0.23	0.21
T/L	6.30	4.74	1.80	2.55	3.00

Table 11 - Morphological characters of *P. angustifolia*, the two subspecies of *P. balsamifera* and intermediate specimens, derived from the correspondence analysis results.
bals. : subsp. *balsamifera*
tri. : subsp. *trichocarpa*

	P. angustifolia	intermediates	*P. balsamifera* (subsp. in brackets)
basal angle : A1	< 55°	50° to 65°	> 60°
width : W/L	< 0.45	0.50 to 0.70	> 0.60
petiole : P/L	< 0.30	0.30 to 0.40	> 0.35
distance to max. width : D/L	> 0.35	0.30 to 0.40	-up to 0.40 (*bals.*) -down to 0.17(*tri.*)
number of teeth : T/L	> 4.5	variable : 1.9 to 6.5	-down to 1.8 (*bals.*) -up to 4.5 (*tri.*)

correspond to subspecies *trichocarpa* (Brayshaw 1965b) : their leaves have short distances to maximum width and large number of teeth.

Between those two extremes, there is a continuous variation of the leaf shape showing multiple intermediate patterns.

- The individuals considered as putative hybrids between *P. angustifolia* and *P. balsamifera* are relatively intermediate.

6 - Intra-individual variability of the morphological characters

Brayshaw (1965b) and Barnes (1969) recommended the collection of leaf material in the median part of the crown. This directive was followed during collection, but it is still important to check the intra-individual variability of the morphological characters. To this effect, short shoots were collected on different parts of a tree. The selected tree was located near Spruce Grove (25km west of Edmonton). Twigs were collected on the north, south, east and west sides of the tree and at three different levels of the crown (high, middle and low). More samples were also collected inside the crown and on basal branches.

As for the general sample, the measurements were made on three leaves for each of two short shoots of each twig. The means were computed for each measurement and transformed as

for the morphological data matrix. The results show a remarkable uniformity for the short shoots collected in the various parts of the crown :

-A1 varies from 70° to 76°

-W/L varies from 0.60 to 0.66

-P/L varies from 0.58 to 0.70

-D/L varies from 0.29 to 0.32

-T/L is about 4

This corresponds to a *P. balsamifera* individual with a pattern intermediate between the two extremes defined in table 11.

However, the short shoots collected on the basal branches of the trunk show a different pattern. W/L is inferior to 0.50 which means much narrower leaves than the leaves collected in the crown of the tree. The petiole is also shorter (P/L < 0.55), the distance to the maximum width larger (D/L 0.34 to 0.38) and the basal angle smaller (A1 50° to 60°). These results confirm the observations of Brayshaw (1965b) and Barnes (1969).

To check the stability of leaf morphology from year to year, collections were made at Kananaskis 4, the first time during the summer of 1986, with the rest of the general sample, and the second time during the summer of 1987. The morphological results again show a great uniformity between the two collections. From year to year, leaves of the short shoots maintain the same pattern on a same tree.

The data of this second collection of Kananaskis 4 were transformed into discrete data and then compared to the first collection data. Values for both collections were in the same classes for 90% of the data (table 12). The second collection data were joined to the general sample data matrix. The results of the correspondence analysis show that the 20 individuals of Kananaskis 4 have approximately the same graph position for both collections.

Table 12 - Morphological data (discrete) for both collections of Kananaskis 4 (1986 and 1987).
The circled values in the second collection indicate the values which are different from the first collection

	First collection (1986)					Second collection (1987)				
	T/L	A1--A2	P/L	W/L	D/L	T/L	A1--A2	P/L	W/L	D/L
K4-01	2	4	3	5	5	2	4	3	5	5
K4-02	2	3	3	4	3	(3)	3	3	4	(4)
K4-03	3	6	6	4	6	3	6	6	4	6
K4-04	3	6	6	4	5	3	6	6	4	5
K4-05	3	5	4	6	4	3	(6)	4	6	4
K4-06	3	6	6	4	5	3	6	6	4	5
K4-07	3	4	4	5	6	3	4	4	5	6
K4-08	2	4	3	5	5	2	4	3	5	5
K4-09	3	7	6	5	7	3	(8)	6	5	7
K4-10	3	7	5	8	6	3	7	5	(7)	6
K4-11	2	2	2	4	3	2	2	2	4	3
K4-12	2	5	5	5	4	2	(4)	5	5	4
K4-13	2	5	3	6	6	2	5	(4)	6	6
K4-14	1	5	4	5	4	2	(6)	(5)	5	4
K4-15	2	5	4	4	4	2	5	4	4	4
K4-16	4	5	4	5	5	4	5	4	5	5
K4-17	3	7	6	6	7	3	7	6	6	7
K4-18	3	6	4	5	6	3	6	4	5	6
K4-19	3	5	3	5	5	3	(6)	3	5	5
K4-20	3	5	4	4	4	3	5	4	(5)	4

V - Flower morphology and pollen study

1 - Flower morphology

Brayshaw (1965a, 1965b) and Vierek and Foote (1970) used flower characters as the only reliable characters to distinguish the two subspecies of *P. balsamifera*.

Twigs bearing flower buds were collected in March 1989 from the trees of the general sample except for the two Waterton locations, which were not accessible at that time of the year. A total of 329 trees at 18 locations was surveyed. The incidence of flowering was low : 160 trees, about half of the trees surveyed, did not bear any flower buds. This flowering rate was similar to that reported by Brayshaw (1965a), who remarked that the proportion of flowering trees varies from one year to the next and even in "good" flowering seasons, not all the trees flower. Possible explanations for this lack of flowering might involve climatic and edaphic conditions, especially harsh in southwestern Alberta, and/or the presence of sterile hybrids in the populations.

The collected twigs, about 30cm long, were placed in water-filled containers in the propagation room of the greenhouse until blooming. Of the 169 samples, 45 could not

be forced to bloom. The remaining 124 samples consisted of 63 females, 59 males, and 2 bisexual individuals.

Female flowers of *P. balsamifera* subsp. *balsamifera* have three carpels, while those of subsp. *trichocarpa* have two carpels. On the 65 female samples examined, the catkins appeared to bear a mixture of 2- and 3-carpellate flowers, which does not allow easy separation into subspecies and validates the lack of specific status for these taxa.

Male flowers of *P. balsamifera* subsp. *balsamifera* have 12 to 20 stamens, while those of subsp. *trichocarpa* have 40 to 60 stamens (Brayshaw 1965a). Among the 30 male individuals of *P. balsamifera* studied, most have an intermediate number of stamens ranging from 20 to 40. Only three individuals have more than 40 stamens. Nine individuals have fewer than 20 stamens. The low flowering rate did not allow the whole population to be tested. However, intermediate number of stamens is frequent in the flowers tested and very few of them have the characteristic numbers of subsp. *trichocarpa*.

The two bisexual individuals were found in the locations Kananaskis 1 (K1-16) and Kananaskis 2 (K2-18). Bisexuality is frequent in *Populus* (Lester 1963, Muhle

Larsen 1970) and reported as deviations of dioecism. These deviations are variable from tree to tree and also from catkin to catkin. The same tree can produce distinct male and female catkins, mixed catkins and hermaphrodite flowers on the same branch. The deviations vary also from one year to the next (Lester 1963) and can disappear completely in certain years (Stettler 1965). Bisexuality is especially frequent in species of section *Leuce*. Lester (1963) reported 38% of deviant trees in populations of *P. tremuloides* from western Connecticut. In balsam poplars, bisexuality was reported less frequently although no general survey was made (Stettler 1965, 1971).

In this study, two individuals exhibited bisexuality of a total of 126 flowering individuals. This represents 1.59% of bisexual individuals among the flowering trees in 1989. The two bisexual individuals had male and female catkins, mixed catkins and a few hermaphrodite flowers on the same branch. Muhle Larsen (1970) and Stettler (1965) reported that interspecific hybrids have a tendancy towards imbalanced sex expression, the amount of deviation from dioecism being significantly higher in hybrids. Lester (1963) and Stettler (1971) concluded that if sex differentiation in *Populus* is controlled by the genome, it is also mediated by auxin levels.

2 - Pollen viability study

Pollen viability has been used in systematic studies to evaluate the reproductive status of putative hybrid individuals (Stebbins 1950). Hybridization can result in the abortive development of anthers, or in irregular meiosis and production of inviable pollen. The resulting proportion of inviable pollen often will depend on how closely related were the two parental species (Grant 1971).

Barnes (1978) studied pollen abortion in *Populus* section *Leuce* to estimate the degree of relationship between species and to identify putative hybrids and backcrosses. Pure species had a pollen viability average greater than 95%. Hybrid F1 pollen viability showed variation between different controlled crosses :

1. - *P. tremuloides* x *P. grandidentata* and *P. grandidentata* x *P. tremula* average 76%. Backcrosses average 71.5%.
2. - *P. tremuloides* x *P. tremula* average 96.5%.

The author concluded that a very close relationship existed between *P. tremuloides* and its European counterpart *P. tremula*.

Eckenwalder (1977) studied pollen abortion in section *Aigeiros* and *Abaso* and found pollen viability in pure species similar to Barnes' results (1978).

To evaluate the pollen viability of the pure taxa in *P. balsamifera*, twigs bearing flower buds were collected in March 1990 from trees located outside the putative hybridization area :

- subsp. *balsamifera* was collected at Plamondon, Lawrence Lake and Smoky River, northeastern Alberta.
- subsp. *trichocarpa* was collected at Burns Lake, Francois Lake and Vanderhoof, western British Columbia.

The twigs were placed in a greenhouse propagation room and male catkins were collected when almost fully elongated, just before anther dehiscence. They were then pressed and dried for further study. Pollen viability was estimated using lactophenol-cotton blue stain (Radford et al. 1974). This stain is highly specific for callose present in mature viable pollen grains, which appear dark blue after macerating for a few hours in the stain solution. Inviable pollen grains are lightly stained or not at all. For each individual, 300 to 400 pollen grains were examined and the proportion of stained pollen calculated.

Estimation of pollen viability in subsp. *balsamifera* ranged from 91% to 98% with an average of 94.7%. In subsp. *trichocarpa*, viability ranged from 91% to 98% with an average of 94.5%.

Collection of *P. angustifolia* samples was not possible but it was assumed that pure *P. angustifolia* would have a similar pollen viability to pure *P. balsamifera*.

The material collected in March 1989 for the flower morphology study was also examined for pollen viability in the general sample. As in most species of *Populus*, species of *Tacamahaca* section are dioecious and only the male individuals can be tested using this technique.

The results (table 13) show a relatively low pollen viability for the 54 individuals tested. Only eight of them exhibit a pollen viability equal to or greater than 94%. The pollen viability in the rest of the sample ranged from 70% to 93% with a few individuals below this mark. Each individual studied was identified using morphological characters described earlier :

1. - Putative hybrids *P. balsamifera* x *P. angustifolia* had a pollen viability ranging from 78% to 92%.
2. - *P. angustifolia* ranged from 63% to 94%.
3. - *P. balsamifera* (both subspecies) ranged from 22% to 99%.

Although the bulk of the population cannot be tested because

Table 13 - Pollen viability of the 54 *Populus* individuals tested; collection March 1989.

Individual	Pollen viability	Morphological Taxon	Remark
CL-01	82%	*balsamifera*	
CR-01	92%	*balsamifera*	
CR-02	94%	*angustifolia*	> or equal to 94%
CR-06	82%	*balsamifera*	
CR-08	78%	*angustifolia*	
CR-10	79%	*putative hybrid*	
CR-13	63%	*angustifolia*	< 70%
CR-19	92%	*putative hybrid*	
DM-01	80%	*angustifolia*	
FC-01	94%	*balsamifera*	>or equal to 94%
FC-02	95%	*balsamifera*	>or equal to 94%
FC-07	96%	*balsamifera*	>or equal to 94%
FC-12	93%	*balsamifera*	
FC-16	73%	*balsamifera*	
FC-19	96%	*balsamifera*	>or equal to 94%
HR-12	76%	*balsamifera*	
HR-14	87%	*balsamifera*	
K1-08	71%	*balsamifera*	
K1-15	47%	*balsamifera*	< 70%
K1-16	75%	*balsamifera*	bisexual
K2-02	80%	*balsamifera*	
K2-03	88%	*balsamifera*	
K2-07	82%	*balsamifera*	
K2-15	84%	*balsamifera*	
K2-18	92%	*balsamifera*	bisexual

Table 13 - continued.

K3-04	22%	balsamifera	< 70%
K3-06	82%	balsamifera	
K3-07	86%	balsamifera	
K3-11	89%	balsamifera	
K4-05	61%	balsamifera	< 70%
K4-13	91%	balsamifera	
K4-17	93%	balsamifera	
SI-02	99%	balsamifera	> or equal to 94%
SI-07	97%	balsamifera	> or equal to 94%
SI-11	92%	angustifolia	
SI-16	93%	angustifolia	
SO-05	88%	angustifolia	
SO-11	71%	angustifolia	
SO-12	84%	angustifolia	
SO-15	91%	putative hybrid	
WC-04	84%	balsamifera	
WC-07	84%	balsamifera	
WC-15	68%	balsamifera	< 70%
WF-04	82%	putative hybrid	
WF-12	81%	balsamifera	
WF-14	76%	balsamifera	
WF-15	61%	balsamifera	< 70%
WF-18	69%	balsamifera	< 70%
WF-19	94%	balsamifera	>or equal to 94%
WF-20	78%	putative hybrid	
WR-04	86%	balsamifera	
WR-05	68%	angustifolia	< 70%
WR-15	75%	angustifolia	
WR-17	66%	angustifolia	< 70%

of dioecism and low flowering rates, these results could indicate the possible occurence of hybridization and subsequent backcrossing betwen *P. angustifolia* and *P. balsamifera* (both subspecies) in the populations studied.

VI - Chemotaxonomy

1 - Definition and history

"Taxonomy is the theory and practice of classification, and chemotaxonomy incorporates the principles and procedures involved in the use of chemical evidence for classification purposes" (Smith 1976).

Chemotaxonomy originated as early as 1741, when De Jussieu observed common medicinal properties in the same group of plants. He suggested "to group plants having similar virtues and also similar forms".
Subsequently, De Candolle (1816) used chemical properties of plants to complement taxonomic studies based on morphological characters. However, it was not until the 20th century that these ideas were fully developed with the advances in chemistry and development of such techniques as chromatography and spectrophotometry. With these techniques, it is possible to isolate and identify chemical constituents present in plants in micro-quantities.

2 - Flavonoids

Not all chemical constituents are useful in a taxonomic study. Primary metabolites occurring in almost all plants, such as peptide amino acids, some fatty acids and sugars, have little or no taxonomic value. The most usefull chemical constituents are usually secondary metabolites such as alkaloids, terpenoids and flavonoids.

In this study, flavonoids have been chosen. This group of compounds presents many advantages :

1. - Constituents of secondary metabolism, such as flavonoids, have a slower turnover than constituents of primary metabolism. Consequently, they will accumulate in the plant and will be easier to detect.
2. - The structural diversity of the flavonoids is important. There are 12 different classes and in each of them, hundreds of compounds have been identified. All of these compounds are widely distributed, mainly in vascular plant families.
3. - Flavonoids are easy to extract from plant material and easy to detect under U.V. light. Furthermore, they are chemically stable.

The taxonomic utility of flavonoids was first demonstrated by Bate-Smith and Lerner (1954) and later by

Bate-Smith (1962). Flavonoids have commonly been used since then to validate or extend existing classifications (Lebreton 1962, Harborne 1966).

Flavonoids can also assist in the study of plant phylogeny. Bate-Smith (1962, 1968) demonstrated the changes in flavonoid biosynthesis which occurred during the evolution of plants. Harborne (1967) defined flavonoid characters categorized as "primitive", "advanced" or "isolated". For example, the presence of leuco-anthocyanins is considered primitive. This group of compounds is present in 60% of the woody plants and absent from 85% of the herbaceous plants. Harborne suggests that the modification from woody to herbaceous habit in the Angiosperms has caused the irreversible loss of the chemical pathways leading to the leuco-anthocyanins. Isolated characters are neither advanced nor primitive but restricted to one plant family or order such as betacyanins in Centrospermae and isoflavones in Leguminosae.

The biological function of flavonoids in plants is poorly understood except for their evident role in floral pigmentation. Various other functions have been assigned to flavonoids, such as growth regulation, as phytoalexins and insect attractors or repellents (Harborne 1973).

Flavonoids are a relatively uniform group structurally; all have the characteristic flavonoid 15 carbon unit (figure 15). Two aromatic rings (A, B) are linked by a three carbon unit (C), which can be open or closed. The three carbon unit

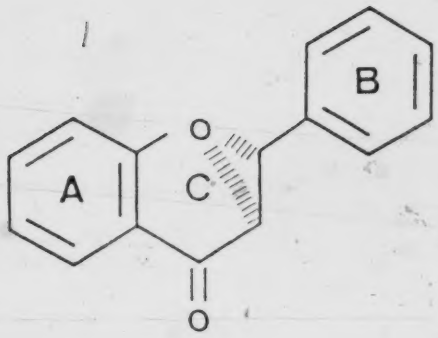


Figure 15 -
Flavonoid 15 carbon unit.

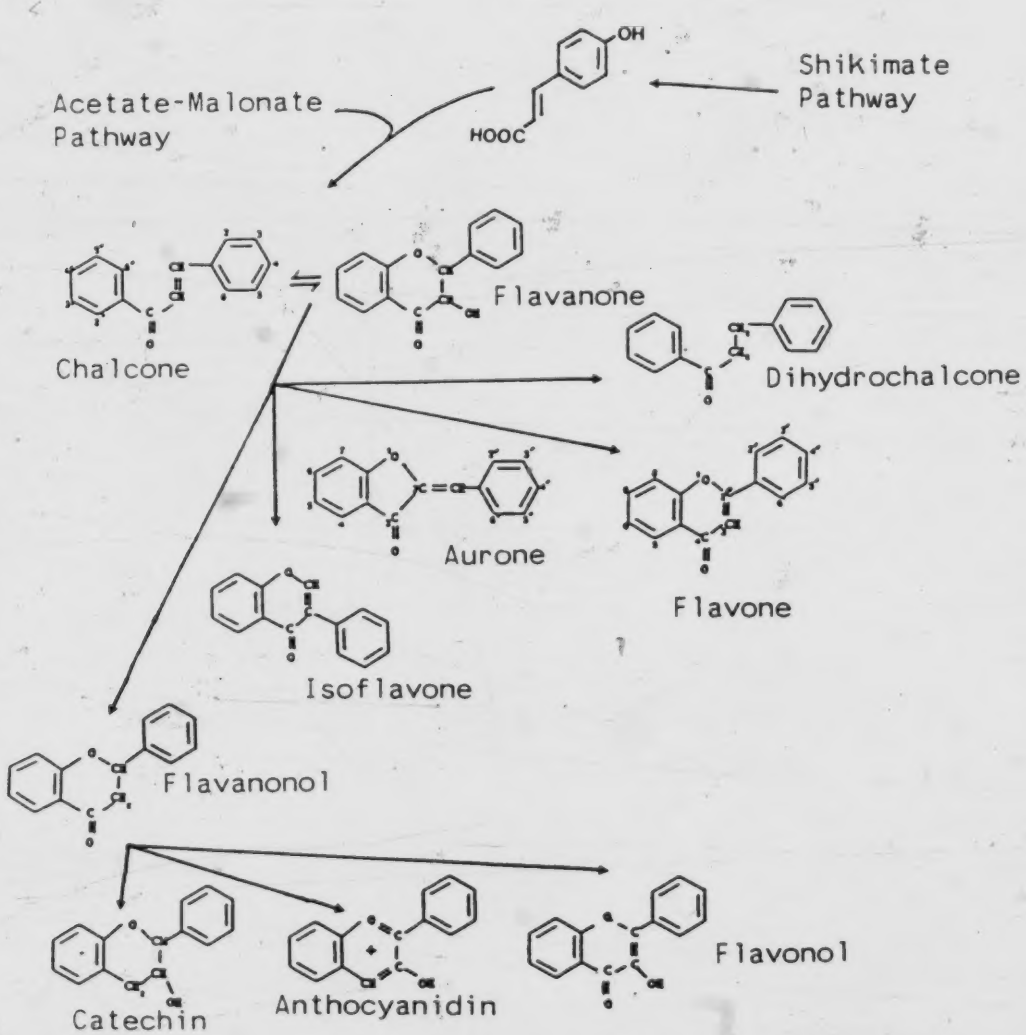


Figure 16 - Flavonoid classes and interrelationships
(Ribéreau-Gayon 1968, Markham 1982)

varies in its level of oxidation to produce 10 different flavonoid classes (figure 16).
Further diversification of the structures is produced by substitutions : hydroxylation, methylation and glycosylation most frequently, but also isoprenylation or esterification. When not combined with a sugar, the flavonoid or aglycone is lipophylic and toxic to the living cell. It is usually located outside the cellular compartment in exudates. Flavonoids are usually found as their glycosides and are hence vacuolar, in the cell solution.

3 - Literature review of the flavonoids in genus *Populus*

Most previous flavonoid studies of *Populus* involved bud exudates. Winter buds in most *Populus* species are coated with a resin, especially abundant in species of section *Tacamahaca*. This resin consists mostly of aglycones mixed with terpenes. Being highly lipophylic, these aglycones are extracted in acetone. Bud exudates have been studied in several species (table 14, part 1), whereas glycosides of the leaves and pollen have been studied in but a few species (table 14, part 2). The flavonoids of most Asiatic poplar species have been totally ignored, as are American species of the section *Tacamahaca*. The leaf glycosides of *P. balsamifera* subsp. *balsamifera* and subsp. *trichocarpa* have

Table 14 - Literature review of the flavonoids in *Populus*.

	Species	Authors
Part 1	Aglycones	
Bud exudates	*P. nigra*	Egger and Tissut (1968) Egger et al. (1969) Chadenson et al. (1971) Wollenweber and Egger (1971) Wollenweber and Weber (1973) Hauteville et al. (1973) Charriere-Ladreix (1973,1976) Tamas at al. (1979)
	P. balsamifera *P. lasiocarpa*	Vanende (1984) Asakawa and Wollenweber (1976) Greenaway et al. (1988)
	all species	Wollenweber (1974,1975)
Part 2	Glycosides	
leaves	*P. candicans*	Thieme (1969)
	P. tremula	Thieme (1966)
	P. x acuminata *P. angustifolia* *P. deltoides*	Crawford (1974) Jones and Siegler (1974)
	section *Aigeiros*	Eckenwalder (1977)
	southwestern North America	
	all species	Hegnauer (1973)
Pollen	*P. yunnanensis*	Sosa and Percheron (1970)

yet to be investigated. In fact, the American balsam poplars, because of their use in papermaking, have been studied for another family of phenolic compounds, the phenol glycosides (Pearl and Darling 1971).

4 - Material and technique

The flavonoid glycosides present in dried leaves were studied. The leaf material was collected during the summer and dried in a plant dryer at 50°C for 48 hours. For each tree sampled, two, two-year-old twigs were utilized. When dried, the leaves were separated from the twigs and petioles, and ground in a coffee grinder.
0.8g of the ground leaves was mixed with 40ml of methanol-water (85-15) in a test tube. The tube was then placed on a shaker for about 16 hours. The slurry was filtered and the ground leaves rinsed with an additional 40 ml of methanol-water (85-15). The methanolic extracts were combined and concentrated using a rotary evaporator, to a minimal volume of about 5ml.
For most samples, the extract was then ready for chromatography, however, some extracts containing excessive amounts of waxes and fats gave poor results. These extracts were washed 3 or 4 times in hexane to remove lipids etc.
Each of the methanolic extracts was examined by two-dimensional chromatography on Whatman 3MM

chromatographic paper. Chromatograms were developed in descending solvent systems :

1. - first dimension : BAW
   Butanol(6) - Acetic Acid(1) - Water(2)
2. - second dimension : AW
   Acetic Acid(1.5) - Water(8.5)

Flavonoids are not generally visible under white light and chromatograms have to be viewed under long-wave ultra-violet light (360nm). The flavonoid zones appear deep purple, light purple, yellow or bright yellow on the chromatogram.
All U.V. zones were then recorded and labelled. The zone colour was noted and the relative flow or Rf. calculated. The Rf. value of a flavonoid is defined as the distance travelled by the compound (from the origin) divided by the distance travelled by the solvent front (from the origin).
The chromatograms were also submitted to different reagents, characteristic of the flavonoids. The different reactions of the flavonoid zones to these reagents can give a first insight into the flavonoid structure :

1. - ammonia fumes (50% in water), some flavonoid purple zones turn yellow or yellow green (Mabry et al. 1970, Markham 1982).
2. - diphenyl-boric acid-ethanolamine complex (Naturstoffreagenz A), spraying of a 1% solution (in methanol) on the chromatogram reveals 3',4'-dihydroxy-flavones and -flavonols as orange zones and 4'-hydroxy- as yellow-green zones.

5 - The synthetic profile

Before studying the flavonoids of the poplar populations of southwestern Alberta, it was necessary to determine flavonoid diversity between populations. During the summer of 1985, a survey was conducted over a broad latitudinal and longitudinal area of Alberta (north limit is High Level, latitude 58°50') and in neighbouring British Columbia. This material was used to study the flavonoid profiles of each species and group of putative hybrids involved in the poplar populations of southwestern Alberta. Supplementary material from Colorado (Gunnison), and British Columbia (Savona and Boston Bar) was also included to complement the information on *P. angustifolia* and *P. balsamifera* subsp. *trichocarpa*.

A synthetic profile for *Populus* sp. was constructed from all 300 chromatograms obtained (figure 17). Table 15 notes the different characteristics of the recorded flavonoid zones.

Each flavonoid zone shown on the synthetic profile was identified. Flavonoids were isolated from about 40 chromatograms for each species and identified using techniques developed by Harborne et al. (1975), Mabry et al. (1970) and Markham (1982). These techniques involve paper and thin layer chromatography (silica and cellulose),

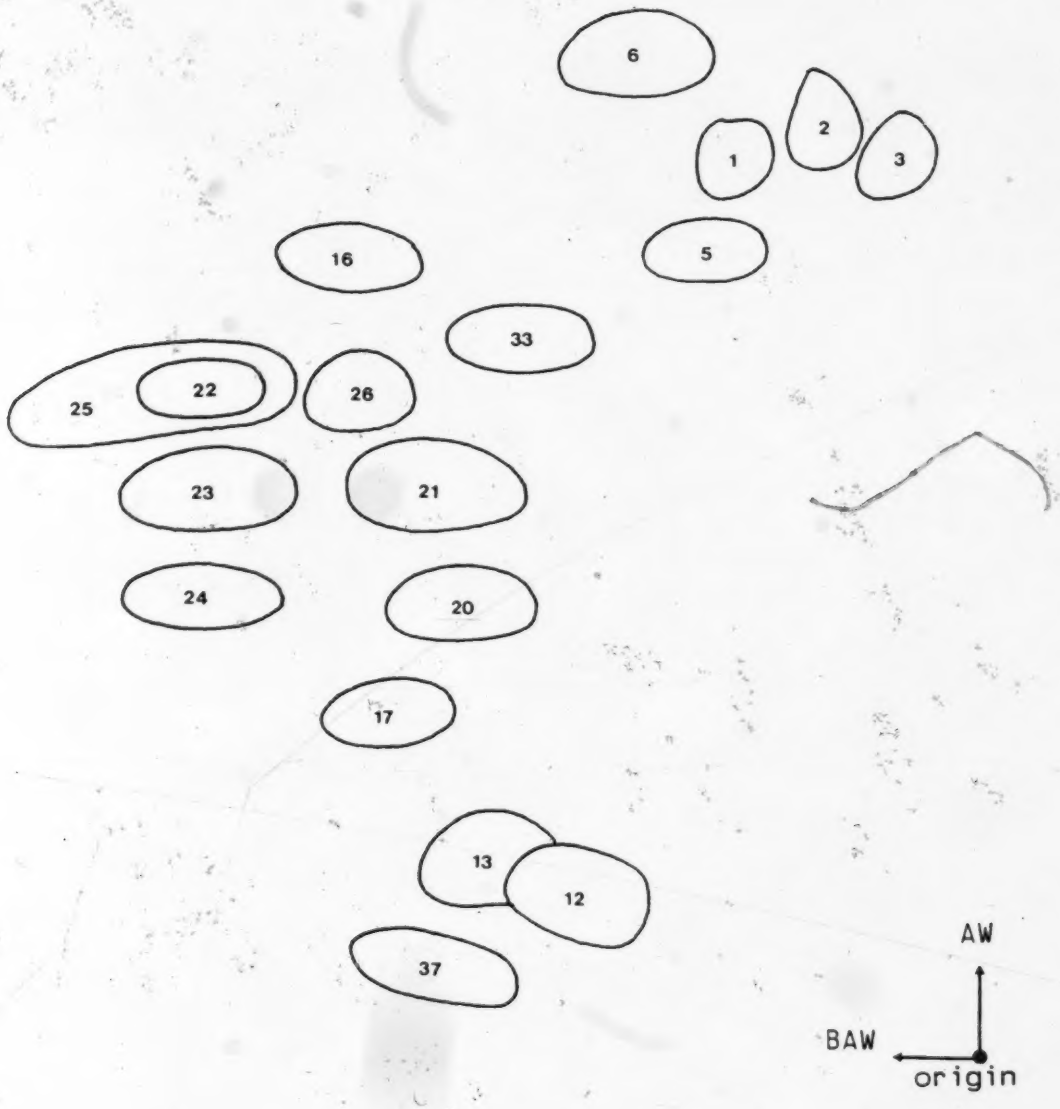


Figure 17 - Synthetic flavonoid profile of the *Populus* taxa studied; each flavonoid zone is delimited and numbered. (chromatographic characteristics in table 15)

- 1st dimension : BAW
  Butanol (60) - Acetic Acid (10) - Water (20)

- 2nd dimension : AW
  Acetic Acid (15) - Water (85)

Table 15 - Chromatographic characteristics of the flavonoid zones recorded on the synthetic profile of the *Populus* taxa studied.
note : Rf. are calculated relative to flavonoid #21

Zone number	Zone colour			Rf.	
	U.V	Ammonia	Na reagent	BAW	AW
1	Purple	Yellow Green	Yellow	0.53	1.50
2	Purple	Yellow Green	Orange Yellow	0.40	1.60
3	Purple	Yellow Green	Orange Yellow	0.32	1.50
5	Purple	Yellow Green	Orange	0.59	1.28
6	Purple	Yellow Green	Yellow	0.75	1.67
12	Purple	Yellow Green	Orange Yellow	0.61	0.48
13	Purple	Bright Yellow	No change	0.78	0.50
16	Grey	No change	No change	1.27	1.37
17	Purple	Yellow Green	Orange	1.00	0.54
20	Purple	Yellow Green	Orange	0.90	0.76
21	Purple	Yellow Green	Orange Yellow	1.00	1.00
22	Purple	Yellow Green	Yellow	1.32	1.18
23	Purple	Yellow Green	Orange Yellow	1.26	0.97
24	Purple	Yellow Green	Orange Yellow	1.42	0.74
25	Grey	No change	No change	1.53	1.11
26	Purple	Yellow Green	Orange Yellow	1.13	1.19
33	Purple	Yellow Green	Orange Yellow	0.91	1.30
37	Purple	Yellow Green	Orange Yellow	0.89	0.33

spectrophotometry, Rf. comparisons and hydrolysis of the flavonoid to identify the sugar(s).

The U.V. absorption maxima of the isolated flavonoids were determined (table 16a), as well as the Rf. of the purified flavonoids (table 16b). These results permitted their determinations (table 16 and figure 18).

The flavonoid glycosides identified belong mainly to the flavonol group, with Kaempferol, Galangin, Myricetin and Quercetin being represented. Three more compounds belong to the flavone group and are based on Luteolin and Chrysoeriol.

---

6 - The species profiles

The profile of *P. angustifolia* was constructed from samples collected in Gunnison, Colorado and Stand-off, Alberta (figure 19a and table 17).

The profile of *P. balsamifera* subsp. *balsamifera* was constructed from samples collected in Plamondon and High Level, Alberta (figure 19b and table 17).

The profile of *P. balsamifera* subsp. *trichocarpa* was constructed from samples collected in Savona and Boston Bar, British Columbia (figure 19c and table 17).

The profile of *P. angustifolia* is characterized by the absence of flavones (#12, 13 and 37), the presence of

Flavonoid stucture	MeOH	NaOMe	NaOAc	NaOAc + Boric Acid		AlCl3+HCl
1	269 294sh 325 356sh	274 406 stable	265 393	266 352	269 354	270 350
2	257 268sh 297sh 356	276 431 stable	262 393	260 379	273 434	269 401
3	255 267sh 296sh 359	272 411 stable	264 395	260 377	270 418	268 365 399
5	259 269sh 296sh 358	264 406 stable	264 392	259 378	274 437	268 402
6	267 297sh 330 363sh	274 405 stable	265 393	266 350	274 351	275 395
12	251 264sh 355	272 421 stable	263 414 decrease	263 381	271 298	266 365 385
13	269 286sh 331	266 403 stable	270 407 decrease	270 339	262 358	258 355
16	266 305sh 344sh	276 390 stable	276 390	266 341	232 273 350	232 273 350
17	253 264sh 304sh 364	266 401 decrease	267 400 decrease	256 368	270 425	270 405
20	256 264sh 304sh 364	267 403 decrease	269 402 decrease	254 370	270 428	271 407
21	258 269sh 300sh 355	280 419 stable	271 318 383	262 374	261 355	263 355
22	266 291sh 306sh 355	266 323 396 stable	273 300 377	266 355	268 305 355	268 309 355
23	257 273sh 296sh 360	275 327 415 stable	271 324 392	261 377	274 434	267 358 400
24	256 267sh 299sh 357	271 332 407 stable	271 325 392	261 377	274 435	268 271 401
25	268 302sh 322sh	278 332 401 stable	274 312 386	267 349	273 302 349	273 302 349
26	257 270sh 296sh 355	280 325 410 stable	271 318 388	261 375	269 432	265 354
33	257 267sh 298sh 357	271 419 stable	271 321 410	258 379	269 423	269 361 401
37	252 268sh 348	258 403 stable	257 398	257 368	272 425	262 357 382

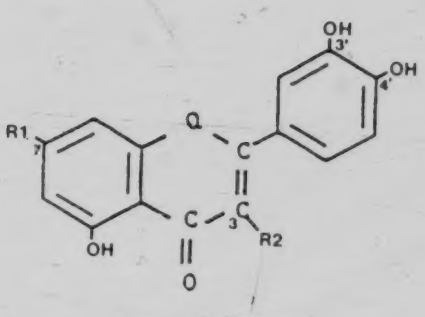
Table 16 - Flavonoid structures isolated from the *Populus* taxa studied.
a - U.V. absorption maxima (nm).
For procedures, see Mabry et al. (1970) and Markham (1982)

Table 16 - Continued.
b - Rf. and probable identity.

flavonoid structure	Rf. BAW	Rf. AW	probable identity
1	.20	.69	Kaempferol 3,7 glucoside galactoside
2	.15	.64	Quercetin 3,7 glucoside galactoside
3	.09	.62	Quercetin 3,7 glucoside rhamnoside
5	.13	.52	Myricetin 3,7 rhamnoside glucoside
6	.23	.72	Kaempferol 3,7 triglucoside
12	.25	.19	Luteolin 7-O glucoside
13	.32	.19	Chrysoeriol 7-O glucoside
16	.49	.67	Galangin 3-MeOH 7-O glucoside
17	.41	.24	Myricetin 3-O glycoside
20	.34	.33	Myricetin 3-O galactoside
21	.39	.46	Quercetin 3-O rhamnoglucoside (rutin)
22	.53	.57	Kaempferol 3-O mono or diglucoside
23	.53	.44	Quercetin 3-O galactoglucoside
24	.54	.29	Quercetin 3-O arabinoside
25	.67	.50	Kaempferol 4'-MeOH 3-O glucoside
26	.45	.52	Quercetin 3-O diglycoside
33	.38	.54	Quercetin 3-O triglucoside
37	.33	.13	Luteolin 7-O galactoside

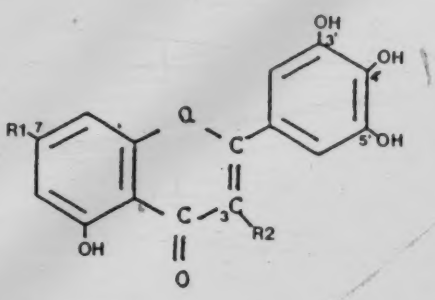
FLAVONOLS

1 - Quercetin



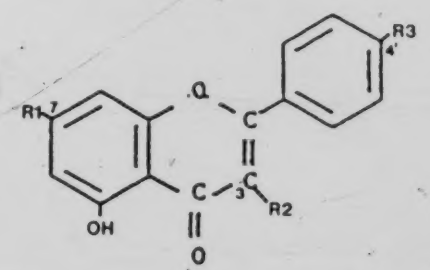
flavonoid # 2 : R1, R2 = glucose, galactose
flavonoid # 3 : R1, R2 = glucose, rhamnose
flavonoid # 21 : R1 = OH, R2 = rutinose
flavonoid # 23 : R1 = OH, R2 = galactose-glucose
flavonoid # 24 : R1 = OH, R2 = arabinose
flavonoid # 26 : R1 = OH, R2 = diglycose
flavonoid # 33 : R1 = OH, R2 = triglucose

2 - Myricetin



flavonoid # 5 : R1, R2 = rhamnose, glucose
flavonoid # 17 : R1 = OH, R2 = glycose
flavonoid # 20 : R1 = OH, R2 = galactose

3 - Kaempferol

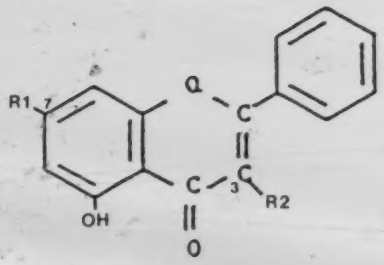


flavonoid # 1 : R1, R2 = glucose, galactose, R3 = OH
flavonoid # 6 : R1, R2 = glucose, R3 = OH
flavonoid # 22 : R1 = OH, R2 = glucose, R3 = OH
flavonoid # 25 : R1 = OH, R2 = glucose, R3 = MeOH

Figure 18 - Identified flavonoid structures of the synthetic flavonoid profile of *Populus* taxa studied.

FLAVONOLS

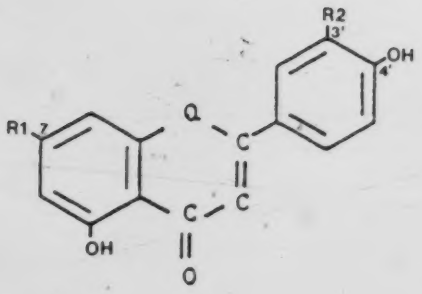
4 - Galangin



flavonoid # 16 : R1 = glucose R2 = MeOH

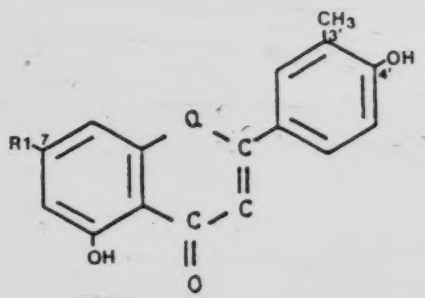
FLAVONES

1 - Luteolin



flavonoid # 12 : R1 = glucose R2 = OH
flavonoid # 37 : R1 = galactose R2 = OH

2 - Chrysoeriol



flavonoid # 13 : R1 = glucose

Figure 18 - continued.

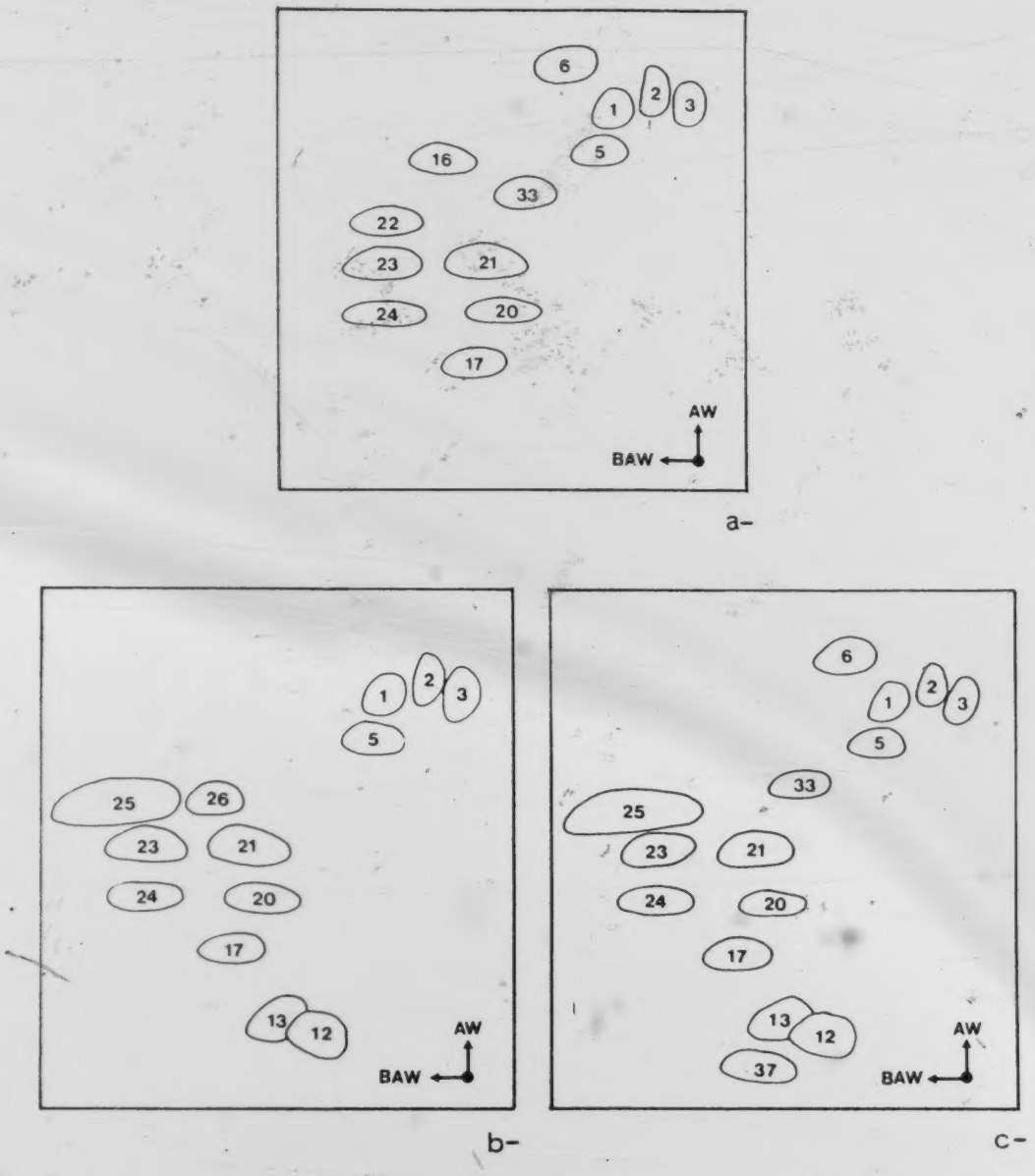


Figure 19 - Flavonoid profiles of the pure *Populus* taxa studied :
a - *P. angustifolia* from Gunnison, Colorado, and Stand-Off, Alberta
b - *P. balsamifera* subsp. *balsamifera* from Plamondon and High Level, Alberta
c - *P. balsamifera* subsp. *trichocarpa* from Boston Bar and Savona, British Columbia.

Table 17 - Presence-absence of the identified flavonoid structures in the flavonoid profiles of *Populus* taxa studied.
a - All structures.

present : X ; absent : 0

Flavonoid stucture	P. angustifolia	P. balsamifera subsp. balsamifera	P. balsamifera subsp. trichocarpa
1	X	X	X
2	X	X	X
3	X	X	X
5	X	X	X
6	X	0	X
12	0	X	X
13	0	X	X
16	X	0	0
17	X	X	X
20	X	X	X
21	X	X	X
22	X	0	0
23	X	X	X
24	X	X	X
25	0	X	X
26	0	X	0
33	X	0	X
37	0	0	X

Table 17 - Continued.
b - Key flavonoid structures for the identification of the taxa studied.

present : X ; absent : 0

taxon 1 : *P. angustifolia*
taxon 2 : *P. balsamifera* subsp. *balsamifera*
taxon 3 : *P. balsamifera* subsp. *trichocarpa*

Flavonoid structure	taxon 1	taxon 2	taxon 3
6 -Kaempferol 3,7 triglucoside	X	0	X
12-Luteolin 7-0 glucoside	0	X	X
13-Chrysoeriol 7-0 glucoside	0	X	X
16-Galangin 3-MeOH 7-0 glucoside	X	0	0
22-Kaempferol 3-0 mono or diglucoside	X	0	0
25-Kaempferol 4'-MeOH 3-0 glucoside	0	X	X
26-Quercetin 3-0 diglycoside	0	X	0
33-Quercetin 3-0 triglucoside	X	0	X
37-Luteolin 7-0 galactoside	0	0	X

Galangin 3-Methyl Ether (#16), and a Kaempferol derivative (#22). In the two subspecies of *P. balsamifera*, flavones are present and Galangin 3-Methyl Ether is replaced by a Kaempferol 4'-Methyl Ether (#25)(table 17).

The distinction is more subtle between the two subspecies of *P. balsamifera*, as there is no demonstrable difference in the aglycone patterns, only in the glycosylation patterns. *P. balsamifera* subsp. *balsamifera* seems to have a less complex glycosylation pattern than *P. balsamifera* subsp. *trichocarpa* (table 17).

In subsp. *balsamifera* :

1. - The Luteolin 7-O Galactoside (#37) is absent.
2. - The highly glycosylated Kaempferol 3,7 (#6) is absent.
3. - The Quercetin 3-O triglucoside (#33) present in the subsp. *trichocarpa* is replaced by a Quercetin 3-O diglycoside (#26).

The concentration of the flavonoids in the extracts was not calculated. However, the extracts were prepared with equal amounts of ground leaves and solvent. It is possible to compare the intensity of the flavonoid zones between chromatograms. When comparing the intensity of the compounds #12 and 13, another difference between the two subspecies of *P. balsamifera* is revealed : the two flavone 7-O glucosides

are dominant in subsp. *trichocarpa*, their colour intensities being very strong; in subspecies *balsamifera*, colour intensities of zones #12 and 13 are weak.

7 - Flavonoid pattern of balsam poplars in southwestern Alberta

Leaf material was collected during the summer of 1986 from the 366 trees of the general sample. Each extract obtained from this leaf material was chromatographied. On the 366 chromatograms so obtained, the presence or absence of each flavonoid zone identified in the synthetic profile was recorded. Correspondence analysis was performed on the data matrix (appendix 3) and the results plotted (figure 20).

The first axis accounts for 48% of the variation in the data. It can be defined as the presence of zones #16 (Absolute Contribution or AC = 0.29) and #22 (AC = 0.33) versus their absence in the flavonoid profile.

The second axis accounts for 21% of the variation in the data. It can be defined as the presence of zone #37 (AC = 0.11) versus the presence of zone #26 (AC = 0.79).

On the plot, several groups can be distinguished. Three of those groups represent flavonoid profiles which can be matched with the species profiles defined earlier :

1. - profiles matching *P. angustifolia* : 71

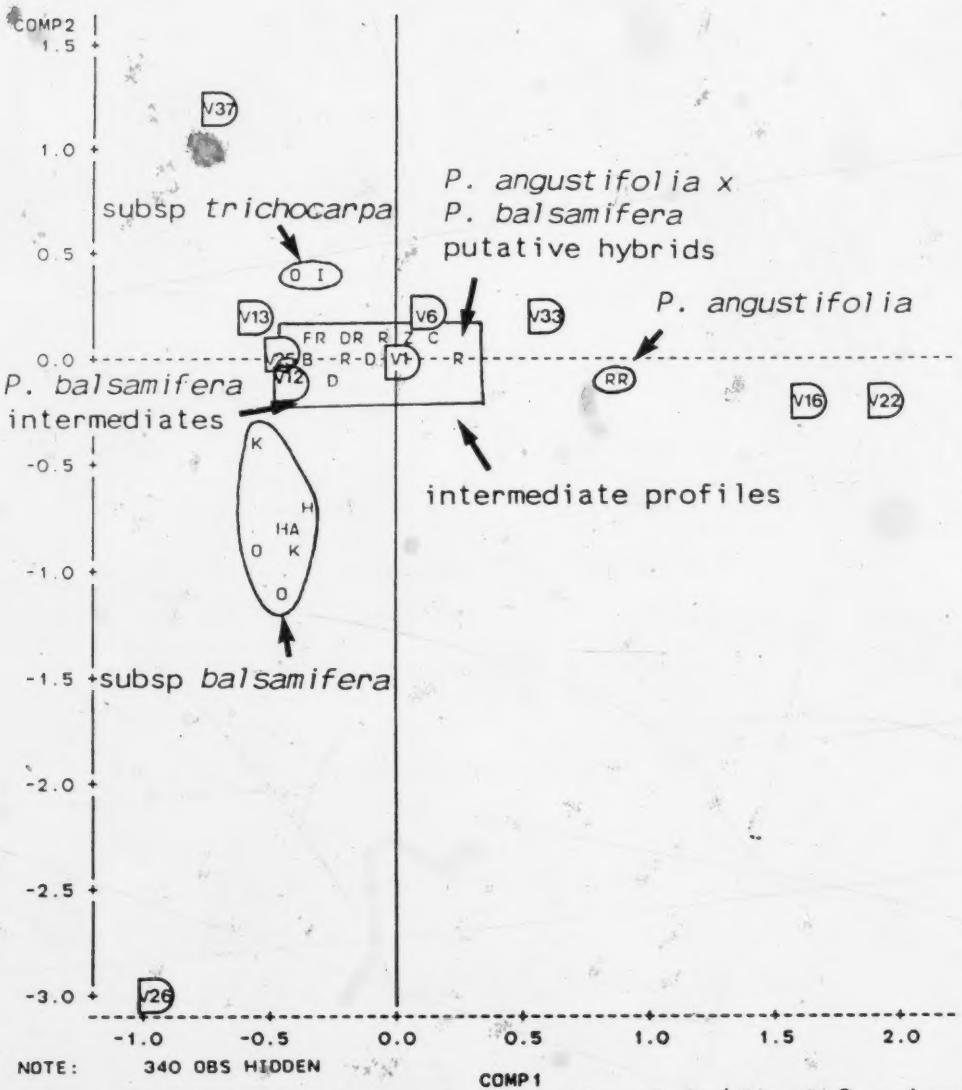


Figure 20 - Correspondence analysis plot (first 2 principal axes) of the flavonoid data matrix.
- Each letter plotted corresponds to an individual.and indicates its location code (location codes are listed in table 5 and chapter III-3).
- flavonoid characters : for example V22 = flavonoid 22

-*P.angustifolia* : 71 individuals
-*P.angustifolia x P. balsamifera* putative hybrids : 20 individuals
-*P. balsamifera* subsp *balsamifera* : 30 individuals
-*P. balsamifera* intermediates : 212 individuals
-*P. balsamifera* subsp *trichocarpa* : 24 individuals

individuals.

These profiles have high positive values (greater than 0.87) on axis 1 and near-to-zero values on axis 2.

2. - profiles matching *P. balsamifera* subsp. *balsamifera* : 30 individuals.

These profiles have negative values on axes 1 and 2.

3. - profiles matching *P. balsamifera* subsp. *trichocarpa* : 24 individuals.

These profiles have negative values on axis 1 and positive values greater than 0.35 on axis 2.

Among these profiles, almost half of them differ slightly from the species profiles by the behaviour of zones #5, 17 and 20. These 3 flavonoids are all myricetin derivatives. One or two of the three can be absent from the profile, in various combinations.

The remainder of the profiles (232 individuals) are intermediate between the species profiles and show great variability. Correspondence analysis was performed on the data matrix including only the flavonoid data of the 232 intermediate profiles. The results show two different groups of intermediate profiles :

1. - profiles intermediate between *P.angustifolia* and *P. balsamifera* : 20 individuals.

They are characterized by the behaviour of flavonoid zones #12, 13 and 16. Three types of intermediate profiles can be described (figure 21a) :

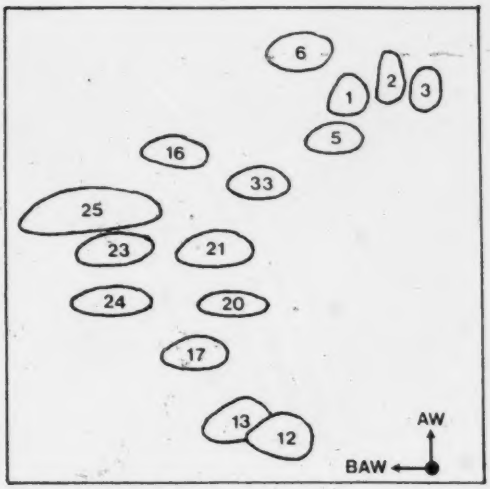
12 13 16 flavonoid zone number

++ ++ -- *P. balsamifera*
-- -- ++ *P. angustifolia*
++ ++ ++ intermediate 1 (5 individuals)
-- -- -- intermediate 2 (4 individuals)
++ -- ++ intermediate 3 (11 individuals)
(++ present, -- absent)

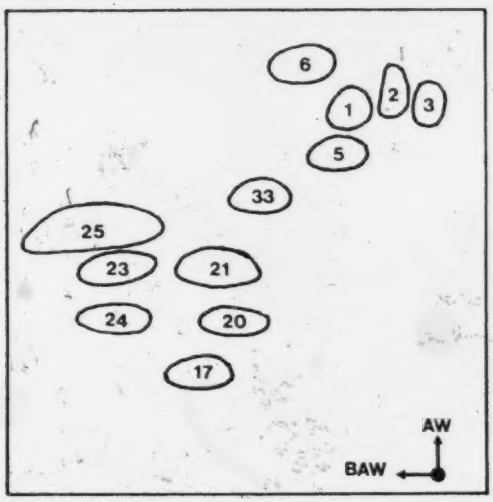
2. - profiles intermediate between subsp. *balsamifera* and subsp. *trichocarpa* : 212 individuals.
They are characterized by the behaviour of flavonoid zone #6, 26, 33 and 37. Some of them have a stronger tendency towards one or the other subspecies profile (figure 21b) :

06 26 33 37 flavonoid zone number

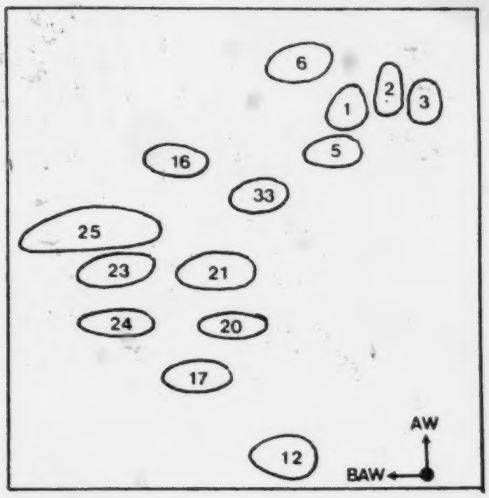
-- ++ -- -- subsp. *balsamifera*
++ -- ++ ++ subsp. *trichocarpa*
++ -- ++ -- intermediate (54 individuals)
++ +- -- -- int. subsp. *balsamifera* (140 ind.)
++ -- -- ++ int. subsp. *trichocarpa* (18 ind.)
(++ present, +- present or absent, -- absent)



Intermediate 1



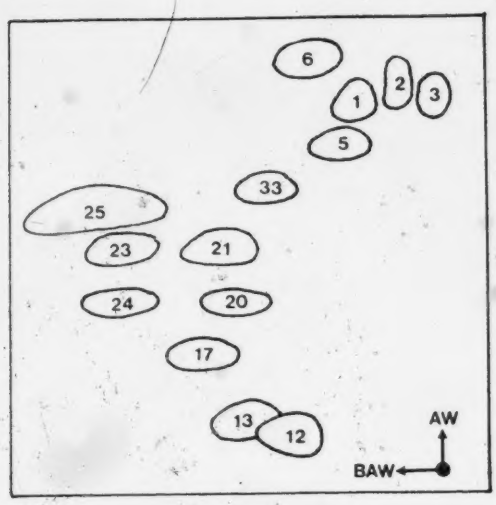
Intermediate 2



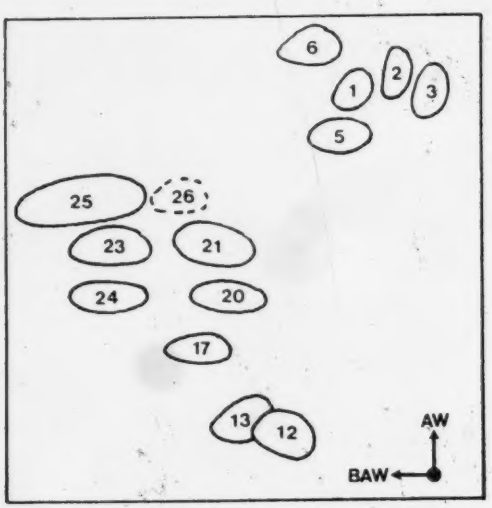
Intermediate 3

Figure 21a - Intermediate flavonoid profiles observed in balsam poplars of southwestern Alberta : 3 types of intermediate profiles between *P. angustifolia* and *P. balsamifera*.

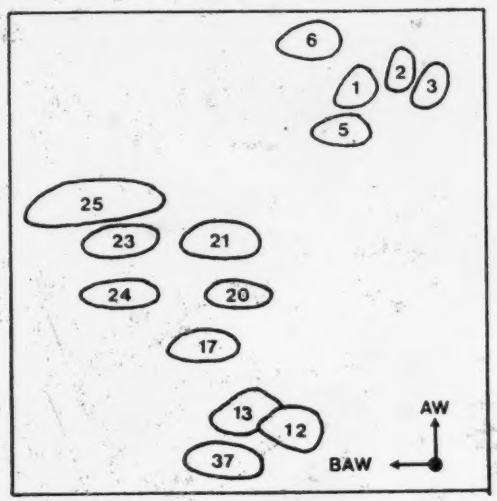
- Intermediate 1 : 5 individuals
- Intermediate 2 : 4 individuals
- Intermediate 3 : 11 individuals



1-



2-



3-

Figure 21b - Intermediate flavonoid profiles observed in balsam poplars of southwestern Alberta :
1 - *P. balsamifera* intermediate, 54 individuals.
2 - *P. balsamifera* intermediate tendency subsp. *balsamifera*, 140 individuals.
3 - *P. balsamifera* intermediate tendency subsp. *trichocarpa*, 18 individuals.

Among the intermediate profiles between subsp. *balsamifera* and subsp. *trichocarpa*, a large number (about 1/3) lack zone #13. This could characterize a tendency of these profiles towards *P. angustifolia*.
For these two groups of intermediate profiles, the flavonoid zones #5, 17 and 20 have the same behaviour as for the profiles matching the species profiles. One or two of the three can be absent in different combinations.

The amount of the different types of flavonoid profiles in each location (table 18) was also studied.
Correspondence analysis was performed on this data matrix and the results plotted (figure 22).
The first axis accounts for 46% of the variation in the data. It can be defined as the presence of *P. angustifolia* (AC = 0.67) versus the presence of *P. balsamifera* subsp. *balsamifera* (AC = 0.14) in the different locations.
Several groups of locations are distinguished on the plot :

1. - Locations composed almost exclusively of *P. angustifolia* have high positive values on axis 1 and low negative values on axis 2. Those locations are Park Lake (PL) and Stand Off (SO).
2. - Locations composed almost exclusively of *P. balsamifera* have negative values on axis 1. One location, Police Outpost (PO), has a majority of subsp. *balsamifera* . Waterton 1 (W1) and Kananaskis 3 (K3) have a large number of subsp.

Table 18 - Distribution of the different types of flavonoid profiles in the 21 *Populus* locations.

A : *P. angustifolia*

B : *P. balsamifera* subsp. *balsamifera*

T : *P. balsamifera* subsp. *trichocarpa*

AB : putative hybrids *P. angustifolia* x *P. balsamifera*

1 : *P. balsamifera* intermediate tendency *balsamifera*

2 : *P. balsamifera* intermediate

3 : *P. balsamifera* intermediate tendency *trichocarpa*

						Intermediates			
	A	B	T		AB		1	2	3
BL	00	00	00	+	00	+	11	05	03
CL	01	00	00	+	03	+	12	03	01
CR	07	00	00	+	03	+	09	01	00
DM	09	00	00	+	00	+	10	00	00
FC	00	00	03	+	00	+	11	04	02
HR	00	02	02	+	00	+	08	08	00
K1	00	00	03	+	00	+	10	07	00
K2	00	02	00	+	00	+	10	06	00
K3	00	05	00	+	00	+	13	02	00
K4	00	02	04	+	00	+	06	04	04
MA	05	00	00	+	00	+	00	00	00
PL	10	00	00	+	00	+	00	00	00
PO	00	13	00	+	00	+	03	01	01
SI	08	00	02	+	02	+	03	03	02
SO	18	00	00	+	01	+	00	00	00
TA	02	00	00	+	00	+	00	01	00
W1	00	04	01	+	00	+	11	04	00
W2	00	00	00	+	00	+	10	00	00
WC	00	02	04	+	01	+	05	03	05
WF	11	00	00	+	03	+	05	01	00
WR	05	00	05	+	06	+	03	01	00

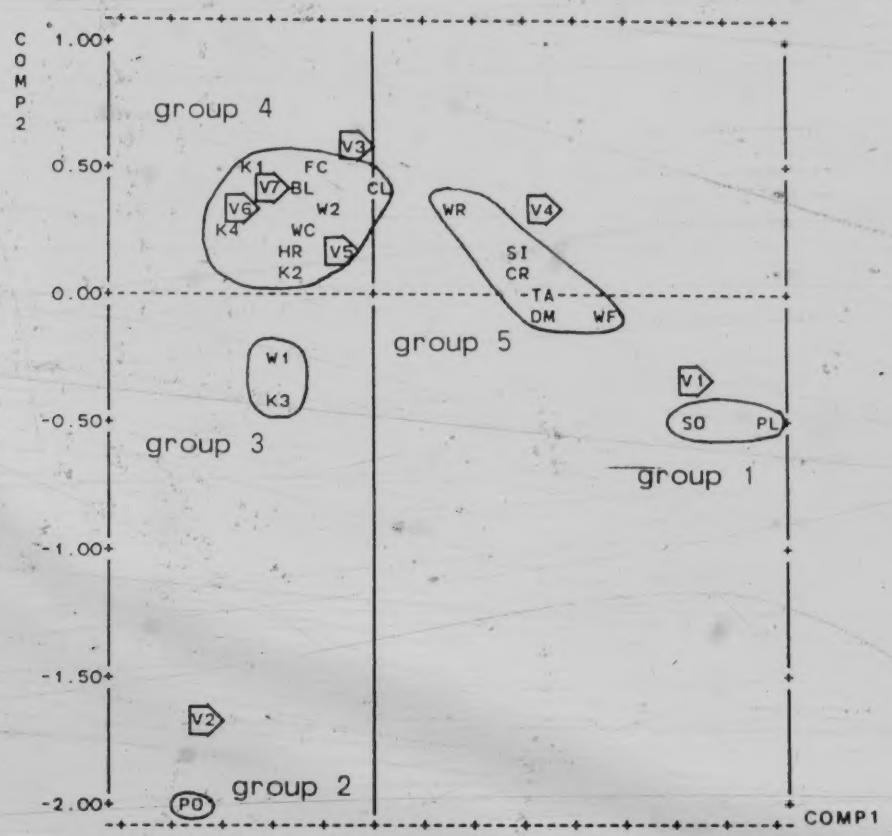


Figure 22 - Correspondence analysis plot (first 2 principal axes) of the flavonoid profile distribution in the 21 *Populus* locations.

- group 1 : *P. angustifolia* locations
- group 2 + group 3 : locations with a large number of *P. balsamifera* subsp. *balsamifera* profiles
- group 4 : *P. balsamifera* intermediate locations
- group 5 : locations with both *P. angustifolia* and *P. balsamifera* plus putative hybrids

V1 : *P. angustifolia* profiles
V2 : *P. balsamifera* subsp. *balsamifera* profiles
V3 : *P. balsamifera* subsp. *trichocarpa* profiles
V4 : *P. angustifolia* x *P. balsamifera* intermediate profiles
V5 : *P. balsamifera* intermediate profiles tendancy subsp. *balsamifera*
V6 : *P. balsamifera* intermediate profiles
V7 : *P. balsamifera* intermediate profiles tendancy subsp. *trichocarpa*

balsamifera profiles and intermediate profiles tending towards subsp. *balsamifera*. The other locations have an intermediate make-up between the two subspecies of *P. balsamifera*. None of the locations has a majority of intermediate profiles tending towards subsp. *trichocarpa* .

3. - Locations having both *P. angustifolia* and *P. balsamifera* plus intermediates have near-to-zero to low positive values on axis 1 and axis 2.

Even if *P. angustifolia* is present as pure individuals in only half of the locations, its influence is still detectable in almost all locations by the suppression of flavonoid zones #12 or 13 in the flavonoid profiles. The only location without this characteristic is Police Outpost. Police Outpost is the southernmost location of the sample and is also the only location to be represented almost exclusively by individuals with flavonoid profiles characteristic of *P. balsamifera* subsp. *balsamifera*. In the other locations, the tendency towards subsp. *balsamifera* is stronger than the tendency towards subsp. *trichocarpa*.

— Crawford (1974) and Jones and Seigler (1974) partially described the *P. angustifolia* flavonoid pattern while studying its putative hybridization with *Populus* species of

the section *Aigeiros*. In a few samples collected in southwestern Alberta, they noted the presence of an additional unidentified zone at the time. The Rf. and other chromatographic data correspond to zone #12, luteolin 7-0 glucoside, which is detected in the present study in profiles intermediate between *P. angustifolia* and *P. balsamifera*. These *P. angustifolia* samples from southwestern Alberta may be putative hybrids between *P. angustifolia* and *P. balsamifera*.

8 - Intra-individual variability of the flavonoid profile

To check for intra-individual variability of flavonoid profiles, leaf samples were collected from different parts of the same tree. The selected tree was located near Spruce Grove (25km West of Edmonton) and possesses a flavonoid profile of the type intermediate between subsp. *balsamifera* and subsp. *trichocarpa* tending towards *balsamifera*. Samples were collected on different sides of the tree (north, south, east and west) and in three different parts of the crown (upper crown, middle branches and basal branches). Two more samples were collected from suckers near the base of the trunk.

None of the 12 samples collected in the crown of the tree showed any variation in flavonoid profile. However, the

flavonoid profiles of the sucker leaf samples, showed a lack of flavonoid zones #17 and 20.

The variability of the flavonoid pattern throughout the year was also investigated. Samples were collected every week from May 12th 1987 to October 29th 1987, from a tree located in Edmonton (Saskatchewan River Valley). None of these samples showed any variation in the flavonoid profile, even the last sample of October, which consisted of senescent leaves.

To check the stability of flavonoid profiles from year to year, collections were made twice at Kananaskis 4 : the first, during the summer of 1986 with the rest of the general sample, and the second during the following summer of 1987. The flavonoid profiles of the 20 trees sampled in this location showed no variation between the two different collections of 1986 and 1987.

VII - Allozyme electrophoresis

1 - Definition and theory

"An allozyme is one of the several forms of an enzyme coded for by different alleles at a locus." (Futuyma 1986)

The most common method of studying allozyme variation is gel electrophoresis. Electrophoresis is designed to separate macromolecules such as proteins, according to mobility differences due to size, shape and net charge. Tissue extracts from different individuals are placed in a porous gel made of starch, polyacrylamide or agarose. A current is applied across the gel and the proteins migrate in response to the electric field.
Enzymes of the same size and shape migrate at a rate determined largely by their net charge, the difference between the number of positively charged amino acids and the number of negatively charged ones. If the different allozymes of an enzyme differ in net charges, they will move at different rates and be separated on the gel.
After electrophoresis, the gel is soaked in a substrate and cofactor solution specific for each enzyme to be studied. The product of the reaction is visualized by a staining reagent revealing a coloured band that indicates the

position of the allozyme. The different allozymes of the same enzyme are identified by differences in the position of the bands on the gel. Different individuals of a population can then be characterized by their allozyme banding patterns, also called zymograms.

Since allozymes are different genetic forms of the same enzyme, they have been used extensively since the late 1960s to study genetic diversity in populations (Lewontin and Hubby 1966, Harris 1969). Indeed, by taking a random sample of enzymes and determining what proportion is genetically variable, it is possible to estimate the proportion of the gene loci of the population that shows variation (Brewer 1970).

Allozymes are also used in systematic studies (Avise 1974, Buth 1984). In this case, allozyme banding patterns are compared to define the degree of similarity or difference between two taxonomic groups, without attempting to determine the genetic basis of the bands.

2 - Literature review of electrophoretic studies in *Populus*

The taxonomic role of electrophoretic data has been demonstrated for many species of *Populus*, mostly in clonal identification problems (table 19, part 1).

Electrophoretic studies of North American balsam poplars have typically involved genetic polymorphism. Weber (1980) and Weber and Stettler (1981) studied 10 populations of *P. trichocarpa* (*P. balsamifera* subsp. *trichocarpa*) from the Pacific Northwest (Washington and Oregon). They found great genetic similarity within and among populations and explain this lack of differentiation by a constant gene flow between these populations, favored by the mating system of the species.
Comtois et al (1986) used electrophoretic data to determine the polyclonal or monoclonal status of *P. balsamifera* stands in the Richmond Gulf (northern Québec).
Other *Populus* species have also been studied for genetic polymorphism (table 19, part 2).

Electrophoresis has also been used to study non enzymatic proteins in *Populus*. Vilar et al. (1988) used electrophoretic data based on stigmatic proteins to

Table 19 - Literature review of electrophoretic studies in *Populus*.
1 - Species and clonal identification studies
2 - Genetic polymorphism studies

	Authors	Species
1	Kim and Chung (1974)	*P. alba*, *P. glandulosa* (Korean aspen), *P. x euramericana*
	Wang et al. (1982)	clones of section *Aigeiros*
	Wang et al. (1983)	genus *Populus*
	Cheliak and Pitel (1984a)	*P. tremuloides* clones
	Fineschi et al. (1986)	clones of *P. x euramericana*, *P. nigra* and *P. deltoides*
	Bergman (1987)	*P. tremuloides* clones
	Castillo and Padro (1987)	*P. x euramericana* clones (I214 and Campeador)
	Ceulemans et al. (1987)	clones of *P. deltoides*, *P. trichocarpa*, *P. maximowiczii* and hybrid clones
	Rajora (1988)	*P. maximowiczii* clones
	Rajora and Zsuffa (1989)	*P. x canadensis* clones (syn. *P. x euramericana*
2	Guzina (1974)	*P. deltoides*
	Rajora (1986)	*P. deltoides*, *P. nigra* and *P. maximowiczii*
	Rajora and Zsuffa (1989)	*P. x canadensis*

differentiate clones of *P. nigra*, *P. alba*, *P. x canescens*, *P. trichocarpa* and *P. x interamericana*.

3 - Material and technique

a - Selection of material

Three types of material have been used to study allozyme variation in *Populus* : young leaves (Comtois et al. 1986), winter buds (Barnes, unpublished) and root tips (Weber 1980, Rajora 1986).

Balsam poplars are characterized by the presence of a resin that is especially abundant on winter buds and on young leaves. A first electrophoretic attempt with these two kinds of tissue, coated with resin, was unsuccessful. The resin prevented the proper migration of the enzymes through the gel, leading to smeared results instead of a clear banding pattern. A second attempt was made after cleaning the resin off the tissue with acetone. Young leaf tissue gave poor results, the enzyme activity being very weak or non-existent. On the other hand, bud tissue gave satisfactory results for nine enzyme systems investigated (table 20).

Root tip tissue does not have the disadvantage of the resin and produced satisfactory enzyme activity for 16 enzymatic systems. However, the results seemed to be non-repeatable. Cuttings were collected on three different trees and rooted

Table 20 - Enzyme systems investigated in buds and roots of *Populus* taxa studied.
- X indicates a successfully investigated enzyme
- XVAR indicates an enzyme for which roots exhibit within-individual variability

Enzyme System	Enzyme Code	Roots	Buds
Alanine Aminopeptidase	AAP	XVAR	0
Aspartate Aminotransferase	AAT	X	X
Alcohol Dehydrogenase	ADH	XVAR	X
Acid Phosphatase	APH	XVAR	0
Esterase	CE	X	X
Glucose-6-Phosphate Dehydrogenase	G6P	X	0
Isocitrate Dehydrogenase	IDH	X	X
Leucine Aminopeptidase	LAP	X	X
Malate Dehydrogenase	MDH	XVAR	X
Malic Enzyme	ME	XVAR	0
Peroxidase	PER	XVAR	0
Phosphoglucose Isomerase	PGI	XVAR	X
Phosphoglucomutase	PGM	XVAR	X
6-Phosphoglucose Isomerase	6PG	XVAR	0
Shikimic Dehydrogenase	SKDH	X	X
Superoxide Dismutase	TO	XVAR	0

in sand, in a propagation room. After three to four weeks, vigorous root tips at the same developmental stage were individually studied by electrophoresis. For each tree, 10 to 15 roots were studied. Ten of the 16 enzyme systems investigated, exhibited variability of the banding pattern between roots of the same tree (table 20).

In spite of the smaller number of identifiable enzyme systems, bud tissue was chosen to study allozyme variation. Indeed, while focusing the different aspects of the technique, approximatively 150 buds from the same tree were processed, and no within-tree variability was detected.

b - Collection of material and preparation

Winter buds were collected at the end of March 1987. On each of the 366 trees in the general sample described earlier, terminal buds were collected from branches of the crown. No difference was found between axillary buds and terminal buds. However, terminal buds were preferred because of their significantly larger size, especially in *P. angustifolia*. For each of the 366 trees, two buds were prepared and analysed individually. Buds were dissected, discarding all the scales and keeping only the green parts, which included embryonic leaves, leaf primordia and the apical meristem (Critchfield 1960). This material was thoroughly washed with acetone and placed in vials containing extraction buffer (table 21). The vials were frozen at -20°C and kept until needed. Storage for more than three months had no effect on

Table 21 - Extraction buffer, gel buffer and electrode buffer recipes used for the electrophoresis of *Populus* samples.

Extraction Buffer (modified from Feret, 1971)

Ascorbic Acid----------0.2g
Cysteine----------0.095g
Tween----------1ml
10% MgCl----------2ml
10% NaCl----------2ml
Sucrose----------17.1g
Mercaptoethanol----------3drops
In 0.1M Trisma Base pH 8.0
Adjust volume to 100ml and pH 7.5

Ridgeway Gel System (Ridgeway et al. 1970)

- Electrode Buffer
Lithium Hydroxide----------0.06M
Boric Acid----------0.30M
pH 8.1

- Gel Buffer
Trisma Base----------0.03M
Citric Acid Anhydrous-------0.005M
pH 8.5

running voltage----------300V
running distance----------3cm and 6.5cm

Table 21 - continued.

Poulik Gel System (Schaal and Anderson 1974)

- Electrode Buffer
Boric Acid----------0.31M
Sodium Hydroxide----------0.063M
pH 8.1

- Gel Buffer
Trisma Base----------0.08M
pH 8.65 (with Citric Acid Anhydrous)

running voltage----------300V
running distance----------4.5cm

Tris-Citrate Gel System (Siciliano and Shaw 1976)

-Electrode Buffer
Trisma Base----------0.13M
Citric Acid----------0.043M
pH 7.0

Gel Buffer
same as Electrode Buffer diluted 1:14.5

running voltage----------200V
running distance----------5.5cm

the banding patterns or the activity of the enzymes studied. On the day the electrophoresis was performed, samples were allowed to thaw, and the bud tissue transferred to another vial containing two or three drops of fresh cold extraction buffer. The sample was then ground using a homogenizer equipped with a teflon head, adding one or two drops of extraction buffer as necessary. Between each step of the procedure, the samples were kept in an ice-bath to reduce denaturation of the enzymes.

c - Electrophoresis technique

The procedure used in this study is horizontal starch gel electrophoresis (for a detailed account of the technique, see Conkle et al. (1982) and Cheliak and Pitel (1984b)). Three different gel systems were utilized (recipes table 21).

The gels were prepared with 12.5% starch (hydrolysed). For one gel, 43.75g of starch and 350ml of gel buffer (table 21) were utilized. About a third of the buffer was used to suspend the starch. The remainder of the buffer was heated to boiling in a microwave oven and then added to the suspended starch. This mixture was heated again to boiling, stirring once or twice, until it became somewhat transparent. Air bubbles were removed under vacuum and the mixture was poured into a gel form (22.5x7.5cm, 0.5 or 0.9mm thick). The gel was then allowed to cool to room temperature, protected from dehydration by a plastic film.

Filter paper wicks (W3, 1x12mm) were placed in the ground tissue samples. The gel was trimmed of excess starch and a cut was made through it about 20mm from the intended cathodal end. The strip so formed was pulled back to allow the loading of the samples. Sample wicks were placed in the cut, about 1mm from each other. Wicks dipped in a standard sample were loaded after every group of 10 sample wicks. This standard was made of radical tissue from germinated corn, ground in the same extraction buffer as the samples. Wicks dipped in red food colouring were placed at the two extremities and at the center of the gel to indicate the migration front. Once so loaded, the two parts of the gel were pushed back together.

The gel was placed on an electrode tray containing the appropriate electrode buffer (table 21). Paper towels (Kimtowels from Kimberley-Clark) were used to conduct the current through the gel. Electrophoresis was performed at 200 or 300 V depending on the gel system (table 21). When the dye migrated 2mm into the gel, the wicks were removed and an ice-pack was placed on the top of the gel. Electrophoresis was continued until the buffer front indicated by the dye had migrated the required distance (table 21).

The gel was sliced by drawing a nylon thread along plastic guides. The upper slice was discarded and the remaining ones (3 to 5 depending on the thickness of the gel) were stained for the various enzyme systems (recipes tables 22 and 23).

Table 22 - Enzyme stain recipes used for the electrophoresis of *Populus* samples.

Ridgeway Gel System 3cm

- Alcohol Dehydrogenase (ADH) E.C.1.1.1.1.
(Shaw and Prasad 1970)
0.2M Trisma Base pH 8.0--------90ml
NAD 1%-------------------------2ml
MTT 1%-------------------------2ml
NBT 1%-------------------------2ml
PMS 1%-------------------------2ml
95% Ethanol--------------------2ml
Incubate at 37°C in dark

- Esterase (CE) E.C.3.1.1.1.
(Collier and Murray 1977)
0.2M Tris-Maleate pH 6.4-------100ml
Fast Blue BB 10%---------------2ml
alpha-Naphthyl Acetate
dissolved in Acetone-----------40mg

- Leucine Aminopeptidase (LAP) E.C.3.4.11.1.
Shaw and Prasad 1970)
0.2M Tris-Maleate pH 5.3-------100ml
MgCl 10%-----------------------2ml
L-Leucyl-beta-Naphthylamide----50mg
Fast Black K Salt--------------70mg
Dissolve 2 substrates in Acetone
Incubate at 37°C in dark

Ridgeway Gel System 6.5cm

Phosphoglucose Isomerase (PGI) E.C.1.1.1.44.
(Yeh and O'Malley 1984)
0.2M Trisma Base pH 8.0--------90ml
NADP 1%------------------------2ml
MgCl 10%-----------------------2ml
MTT 1%-------------------------2ml
PMS 1%-------------------------2ml
Glucose-6-P Dehydrogenase------2ml
Fructose-6-Phosphate-----------50mg
Incubate at 37°C in dark

Table 22 - continued.

Poulik Gel System

Aspartate Aminotransferase (AAT) E.C.2.0.1.1.
(Siciliano and Shaw 1976)
0.2M Trisma Base pH 8.0--------50ml
Fast Blue BB 10%---------------2ml
AAT substrate solution---------50ml
Incubate at 37°C in dark

Tris-Citrate Gel System

- Isocitrate Dehydrogenase (IDH) E.C.1.1.1.41.
(Yeh and O'Malley 1980)
0.2M Trisma Base pH 8.0--------90ml
NADP 1%------------------------2ml
MgCl 10%-----------------------2ml
MTT 1%-------------------------2ml
PMS 1%-------------------------2ml
DL-Isocitric Acid--------------400mg
Incubate at 37°C in dark

- Malate Dehydrogenase (MDH) E.C.1.1.1.37.
(Yeh and O'Malley 1980)
0.2M Trisma Base pH 8.0-----------50ml
DL-Malic Acid substrate solution--50ml
NAD 1%-----------------------------4ml
NBT 1%-----------------------------2ml
PMS 1%-----------------------------2ml
MTT 1%-----------------------------2ml
Incubate at 37°C in dark

Table 22 - continued.

Phosphoglucomutase (PGM) E.C.2.7.5.1.
(Yeh and O'Malley 1980)
0.2M Trisma Base pH 8.0--------92ml
NADP 1%----------2ml
MgCl 10%----------2ml
MTT 1%----------2ml
PMS 1%----------2ml
Glucose-6-P dehydrogenase------2ml
Glucose-1,6 diphosphate--------1ml
Glucose-1-Phospate----------600mg
Incubate at 37°C in dark

Shikimic Dehydrogenase (SKDH) E.C.1.15.1.1.
(Conkle et al 1982)
0.2M Trisma Base pH 8.0--------80ml
Agar----------600mg
Shikimic Acid----------20mg
NADP 1%----------2ml
PMS 1%----------2ml
MTT 1%----------2ml
Dissolve Agar in 40ml of 0.2M Trisma Base. Bring to boil and add remaining 0.2M trisma Base. Stir until slightly cooled and add substrate and dyes. Incubate at 37°C in dark

Abbreviations used for chemical names :
MTT : Thiazolyl Blue
PMS : Phenazine Methosulfate
NBT : Nitro Blue Tetrazolium
NAD : beta-Nicotinamide Adenine Dinucleotide
NADP : beta-Nicotinamide Adenine Dinucleotide Phosphate

Table 23 - Substrate solutions used in enzyme stain recipes for the electrophoresis of *Populus* samples.

- AAT substrate solution
alpha-Ketoglutaric Acid----------1.46g
L-Aspartic Acid----------5.32g
PVP-40T----------20g
EDTA----------2g
Sodium Phosphate Dibasic----------56.80g
Make up to 2L in distilled water

- DL-Malic Acide substrate solution
DL-Malic Acid----------0.5M
Trisma Base----------0.2M
pH 7.0 (with NaOH)

- Glucose-1,6 Diphosphate solution
alpha-D-Glucose-1,6 Diphosphate--------10mg
dissolved in 100ml distilled water

- Glucose-6-P Dehydrogenase solution
Citric Acid Trisodium Salt 0.005M------100ml
Glucose-6-P Dehydrogenase----------5000units

Abreviations :
EDTA : Disodium Salt of Ethylene Diamine Tetraacetic Acid
PVP-40T : Polyvinylpyrrolidone, Molecular Weight 40.000

d - Zymogram patterns

For each enzyme system studied, an attempt to ascribe the different banding patterns to genotypes was made (figure 23). Loci were numbered from the anodal to the cathodal side of the gel. For each locus, the most common allele was named A, the second commonest B, etc. The Rf. for each allele was calculated in comparison with the best band observed on the corn standard banding pattern (table 24).

In the case of the second locus of IDH, the resolution of the banding pattern was not totally satisfactory. The bands presented a well defined front but a smeary tail (figure 23). For this locus, only the front of the band was considered, scoring the fast front as homozygote AA and the slow front as homozygote BB. The presence of the heterozygote AB was not possible to detect.

Observed genotypes for the nine enzyme systems studied were recorded for the 366 trees of the general sample. These data were transformed into discrete data before performing correspondence analysis (appendix 4, includes transformation key).

4 - Results

a - Percentage distribution of each genotype in the general sample

Among the enzyme systems studied, ADH, LAP and SKDH showed

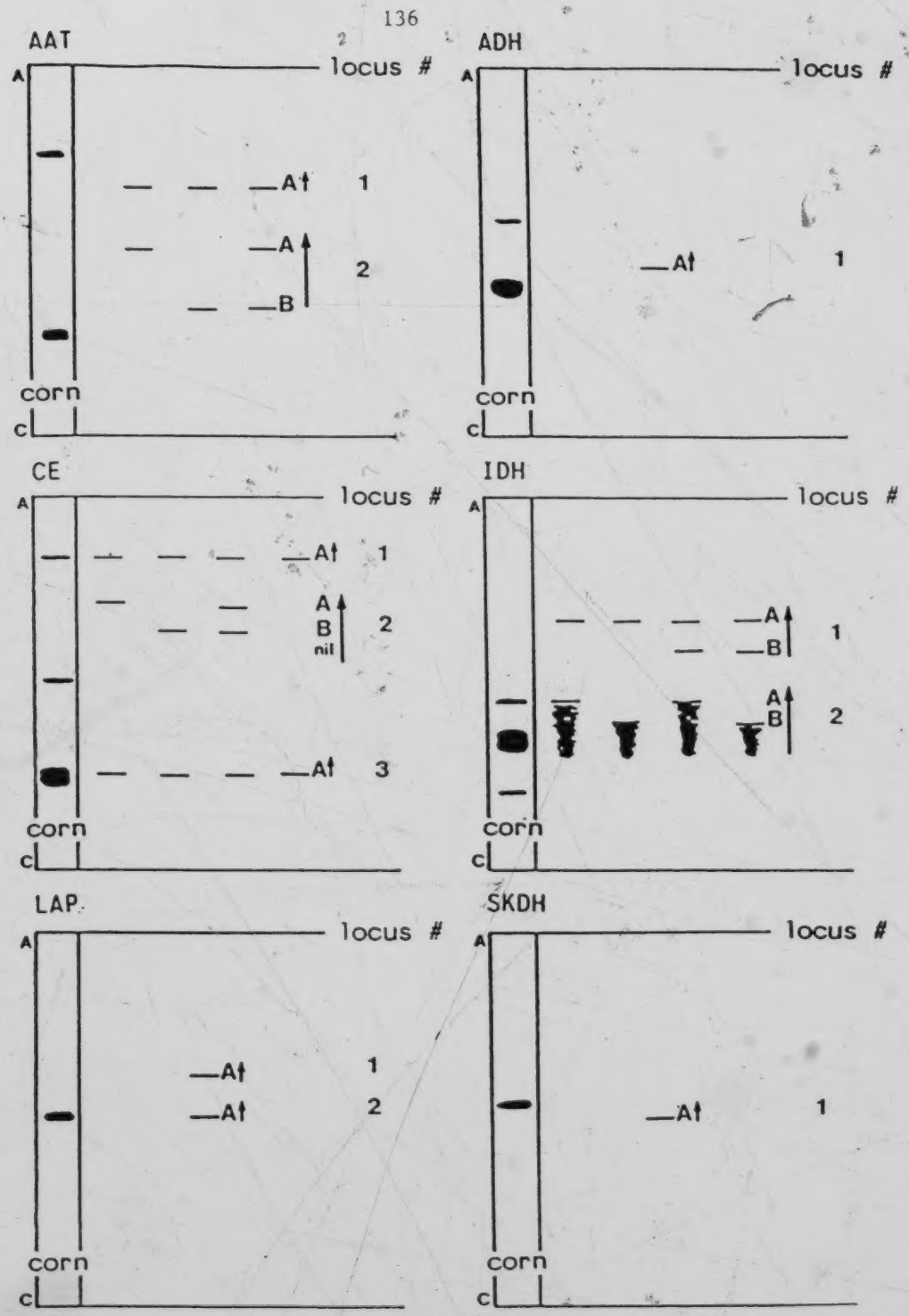


Figure 23 - Zymogram patterns observed in the *Populus* taxa studied (winter buds).

(A = Anode, C = Cathode)

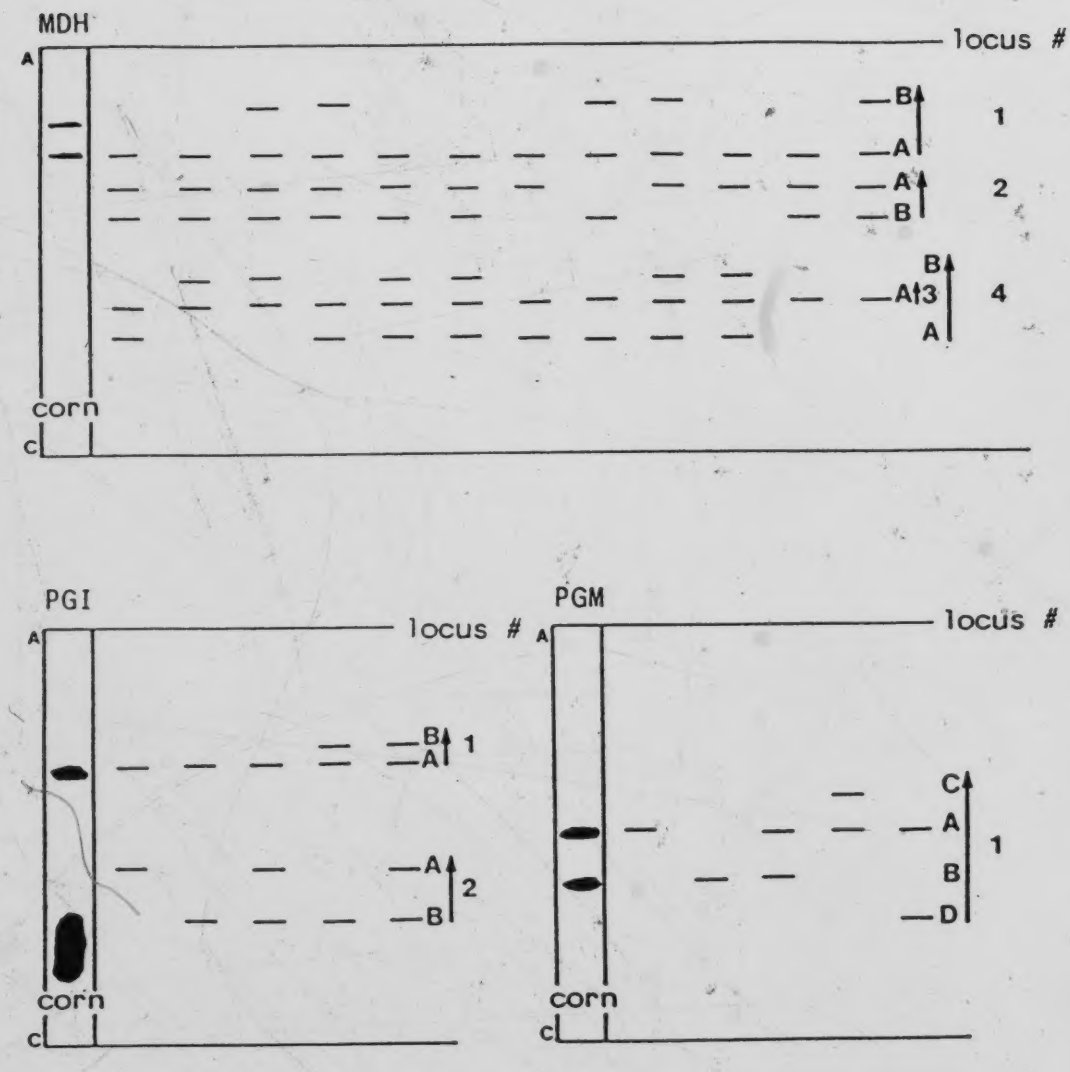


Figure 23 - continued.

Table 24 - Loci and alleles observed for the nine enzyme systems studied in the *Populus* samples.

Enzyme	Locus	Allele	Rf.	Scored against
AAT	1	A	0.83	fastest corn band
	2	A	0.70	
		B	0.64	
ADH	1	A	1.25	bottom of lowest corn band
CE	1	A	1.00	fastest corn band
	2	A	0.92	
		B	0.83	
	3	A	0.43	
IDH	1	A	1.47	fastest corn band
		B	1.29	
	2	A	1.00	
		B	0.84	
LAP	1	A	1.10	corn band
	2	A	1.00	
PGI	1	A	1.00	top of fastest corn band
		B	1.10	
	2	A	0.67	
		B	0.53	
PGM	1	A	1.00	top of fastest corn band
		B	0.86	
		C	1.09	
		D	0.76	
SKDH	1	A	0.77	top of corn band
MDH	1	A	1.00	slowest corn band
		B	1.05	
		C	1.10	
	2	A	0.88	
		B	0.83	
	3	A	0.54	
	4	A	0.41	
		B	0.66	

no variation, as well as the loci AAT1, CE1, CE3 and MDH3 (table 25).

Some loci exhibited rare alleles :

1. - IDH1 varied only for two individuals of the general sample, K3-02 and HR-03.
2. - PGI1 varied only in the Daisy May location for six of the 19 individuals.
3. - PGM1 showed a fast allele C in HR-08 and a slow allele D in HR-09.
4. - MDH showed combinations of alleles represented only in one or a few individuals of the general sample.

b - Correspondence analysis

In correspondence analysis, the first axis accounted for 28% of the variation in the data(figure 24). This axis can be defined as the locus AAT2 (Absolute Contribution or AC=0.38) versus the locus CE2 (AC=0.49) : individuals heterozygous (AB) at AAT2 and homozygous (AA) at CE2 have high positive values on axis 1; individuals homozygous (AA) at AAT2 and not showing the locus CE2 have high negative values on axis 1.

The second axis accounted for 22% of the variation in the data. It can be defined as the loci CE2 (AC=0.18) and PGM1 (AC=0.48) combined, versus the locus MDH4 (AC=0.26) : individuals heterozygous (AB) at PGM1, not showing the locus CE2 and heterozygous (AB) at MDH4 have high positive values

Table 25 - Percentage distribution of each genotype in the population of *Populus* taxa studied.

Locus	genotype	%	Locus	genotype	%
AAT1	AA	100%	MDH1	AA	76.37%
				AB	23.63%
AAT2	AA	62.36%			
	BB	9.61%	MDH2	AA	4.94%
	AB	28.02%		BB	0.27%
				AB	94.78%
ADH1	AA	100%	MDH3	AA	100%
			MDH4	AA	63.74%
CE1	AA	100%		BB	7.42%
				AB	27.75%
CE2	AA	28.84%		nil	1.10%
	BB	17.31%			
	AB	28.02%	PGI1	AA	98.35%
	nil	25.82%		AB	1.65%
CE3	AA	100%	PGI2	AA	1.92%
				BB	16.76%
IDH1	AA	99.45%			
	AB	0.55%	PGM1	AA	38.19%
IDH2	AA	70.88%		BB	12.36%
	BB	29.12%		AB	48.90%
				AC	0.27%
LAP1	AA	100%		AD	0.27%
LAP2	AA	100%	SKDH1	AA	100%

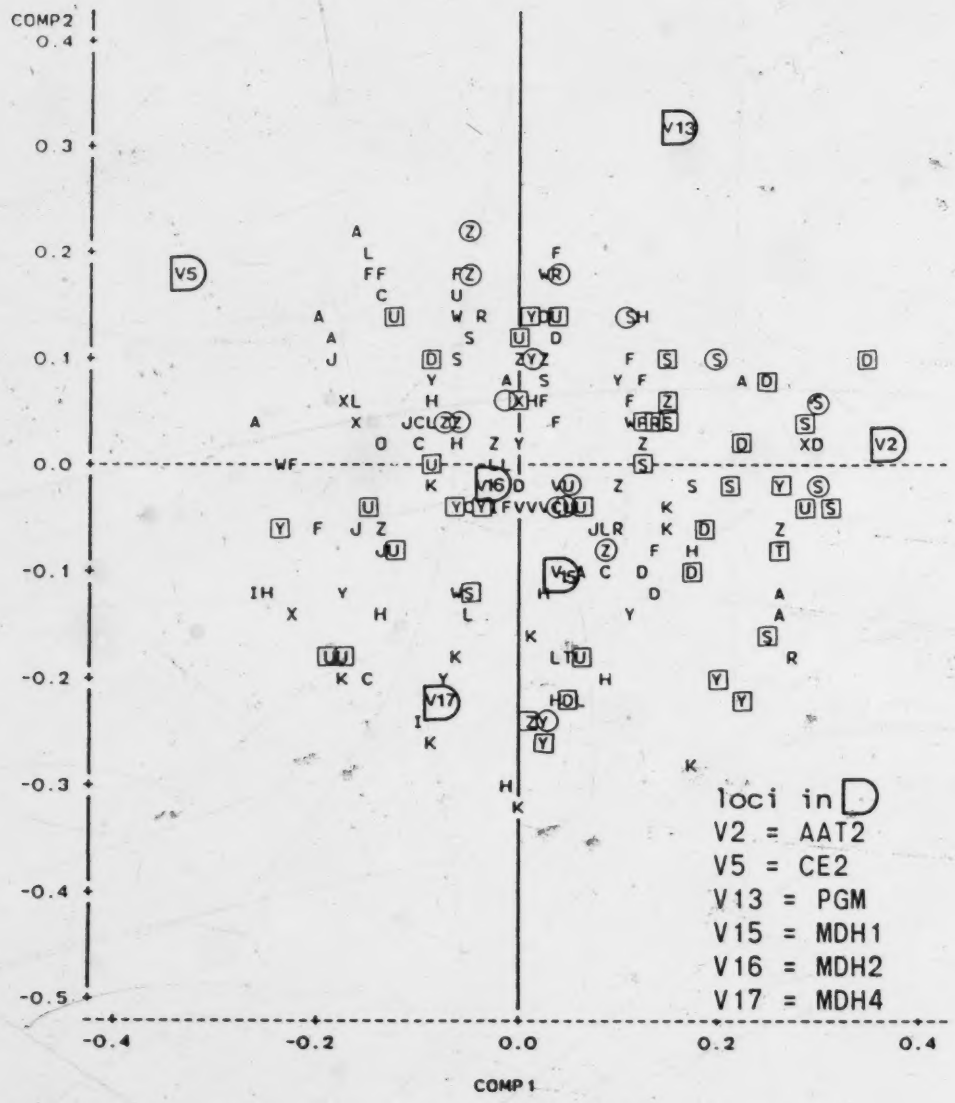


Figure 24 - Correspondence analysis plot (first 2 principal axes) of the electrophoresis data matrix.

-Each letter plotted corresponds to an individual and indicates its location code (location codes are listed in table 5 and chapter III-3).

- *P. angustifolia* individuals are in squares
- *P. angustifolia x P. balsamifera* putative hybrids are circled
- The other individuals are all *P. balsamifera*

on axis 2; individuals homozygous (AA) at CE2, and homozygous (AA) at MDH4 have high negative values on axis 2. The individuals were identified to taxon on the basis of Brayshaw's morphological criteria (1965a) as in the precedent methods (figure 24). There was no clear disjunction between the three groups of species and putative hybrids. However, individuals belonging to *P. angustifolia* and to the group of putative hybrids between *P. angustifolia* and *P. balsamifera* seem to be more densely distributed on the right-hand side of the plot (high positive values on axis 1). The individuals of these two groups are more frequently heterozygous (AB) at AAT2 and homozygous (AA) at CE2.

Individuals belonging to *P. balsamifera* (two subspecies not distinguished) are distributed almost homogeneously throughout the graph.

These results are confirmed when checking the enzyme activity of bud samples from individuals of pure species collected outside the putative hybridization area. No typical species pattern could be identified.

c - Frequency distribution of each genotype in the locations

For the nine enzyme systems studied, the frequency of each enzyme genotype at each location was recorded (table 26). Correspondence analysis was performed on this data matrix (figure 25).

	AAT2			CE2				IDH1		IDH2		PGI1		PGI2			PGM1					MDH1		MDH2			MDH4			loci
	AA	BB	AB	AA	BB	AB	n.	AA	AB	AA	BB	AA	AB	AA	BB	AB	AA	BB	AB	AC	AD	AA	AB	AA	BB	AB	n.	AA	BB	AB	genotypes
BL	19	00	00	00	06	03	10	19	00	16	03	19	00	00	00	19	04	00	15	00	00	19	00	00	00	19	00	19	00	00	
CL	16	01	03	03	02	10	05	20	00	12	08	20	00	00	00	20	07	03	10	00	00	10	10	00	01	19	00	02	00	18	
CR	04	01	15	08	00	03	09	20	00	04	16	20	00	04	01	15	03	02	15	00	00	18	02	00	00	20	00	20	00	00	
DM	03	07	09	10	04	04	01	20	00	13	06	13	06	00	04	15	06	05	08	00	00	10	09	05	00	14	01	14	01	03	
FC	10	04	06	00	09	02	09	19	01	14	06	20	00	00	03	17	04	01	15	00	00	16	04	00	00	20	00	18	00	02	
HR	12	06	02	03	05	02	10	19	01	11	09	20	00	00	00	20	06	04	08	01	01	11	09	00	00	20	00	09	00	11	
K1	20	00	00	00	04	14	02	20	00	20	00	20	00	00	00	20	19	00	01	00	00	20	00	00	00	20	00	00	00	20	
K2	20	00	00	03	00	10	07	20	00	19	01	20	00	00	00	20	05	05	10	00	00	12	08	00	00	20	00	20	00	00	
K3	11	01	08	03	11	06	00	20	00	15	05	20	00	00	00	20	13	00	07	00	00	02	18	00	00	20	00	00	00	20	
K4	20	00	00	06	05	04	05	20	00	12	08	20	00	00	14	06	12	05	03	00	00	15	05	00	00	20	00	18	00	02	
MA	00	00	05	05	00	00	00	05	00	05	00	05	00	00	05	00	05	00	00	00	00	05	00	00	00	05	00	05	00	00	
PL	10	00	00	10	00	00	00	10	00	10	00	10	00	00	10	00	00	00	10	00	00	10	00	00	00	10	00	10	00	00	
PO	07	08	03	00	01	14	03	18	00	18	00	18	00	00	00	18	11	00	07	00	00	18	00	00	00	18	00	18	00	00	
SI	09	01	10	11	05	04	00	20	00	14	06	20	00	03	03	14	03	04	13	00	00	16	04	00	00	20	03	16	01	00	
SO	11	01	08	06	00	12	02	20	00	10	10	20	00	00	11	09	11	03	06	00	00	19	01	00	00	20	00	16	00	04	
TA	01	00	02	03	00	00	00	03	00	02	01	03	00	00	02	01	02	01	00	00	00	02	01	00	00	03	00	03	00	00	
W1	13	02	05	05	02	06	07	20	00	18	02	20	00	00	00	20	06	06	08	00	00	17	03	10	00	10	00	14	00	06	
W2	06	00	04	03	01	01	05	10	00	10	00	10	00	00	00	10	02	00	08	00	00	10	00	00	00	10	00	10	00	00	
WC	10	01	09	06	04	00	10	20	00	14	06	20	00	00	00	20	03	01	16	00	00	13	07	00	00	20	00	05	00	15	
WF	11	00	09	08	03	04	05	20	00	16	04	20	00	00	04	16	14	00	06	00	00	18	02	00	00	20	00	00	20	00	
WR	14	02	04	12	01	03	04	20	00	05	15	20	00	00	04	16	03	05	12	00	00	17	03	03	00	17	00	15	05	00	
locations																															

Table 26 - Frequency distribution of each genotype in the 21 *Populus* locations (only for loci exhibiting variation).

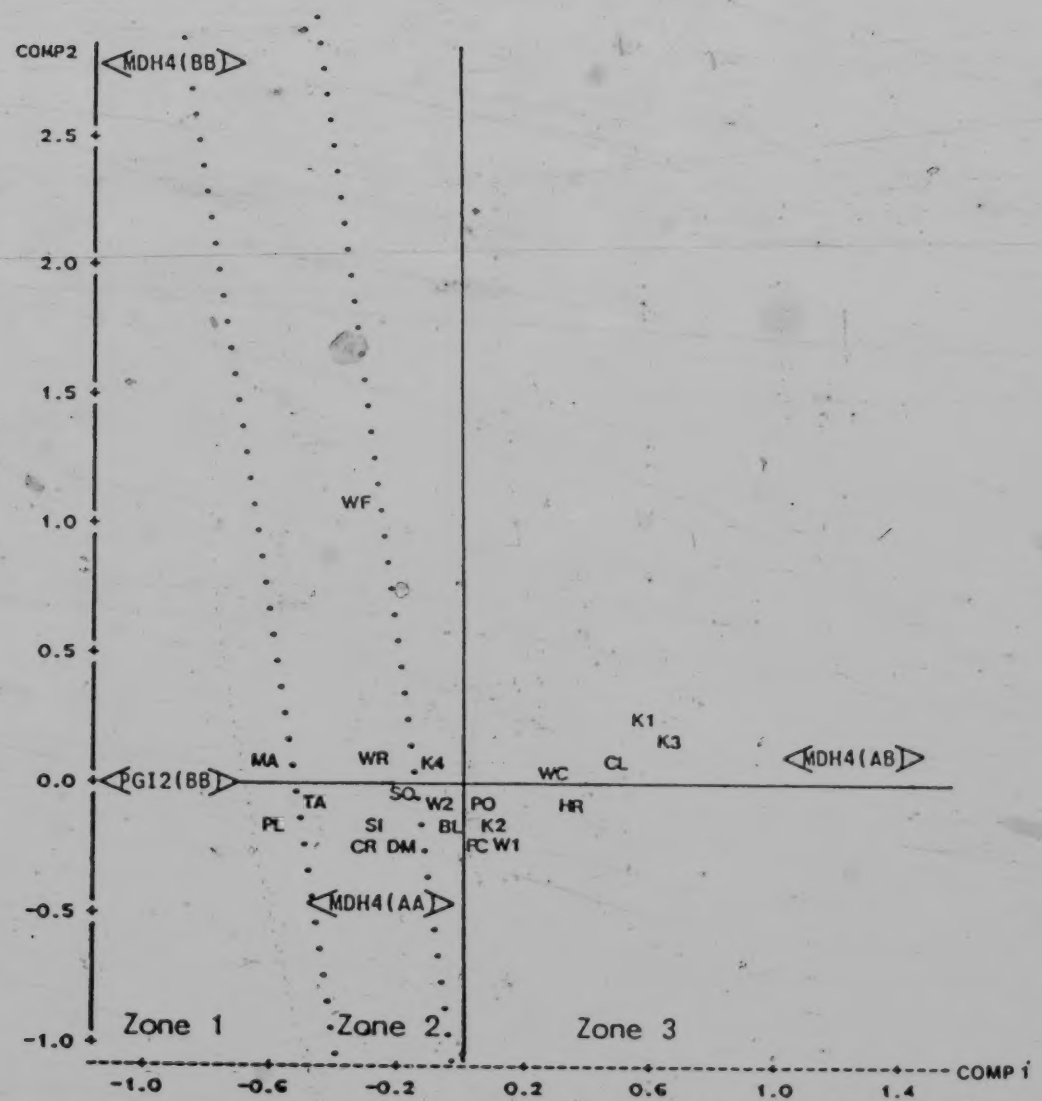


Figure 25 - Correspondence analysis plot (first 2 principal axes) on the frequency distribution data matrix (frequency of each genotype in the 21 *Populus* locations).

- zone 1 : pure *P. angustifolia* locations
- zone 2 : mixed *P. angustifolia* and *P. balsamifera* locations
- zone 3 : pure *P. balsamifera* locations

note : only the significant characters (in < >) with high absolute contributions are plotted, MDH4(AA), MDH4(BB), MDH4(AB) and PGI2(BB)

The first axis accounts for 22% of the variation in the data. It can be defined as the presence of the genotype AB for the locus MDH4 (AC=0.34) versus the presence of the genotype BB for PGI2 (AC=0.15).

The second axis accounts for 16% of the variation in the data. It can be defined as the presence of the genotype BB at MDH4 (AC=0.71) versus the presence of the genotype AA at MDH4 (AC=0.11).

On the plot, it is possible to distinguish three zones :

- in zone 1 are homogenous locations represented only by *P. angustifolia*;
- in zone 2 are mixed locations represented by *P. angustifolia*, *P. balsamifera* and putative hybrids;
- in zone 3 are homogeneous locations represented only by *P. balsamifera* (both subspecies).

P. balsamifera locations are characterized by a higher frequency of the genotype AB for the locus MDH4. *P. angustifolia* locations are characterized by a higher frequency of the genotype BB for the locus PGI2. Mixed locations are intermediate.

d - Synthesis

The different taxa were not clearly differentiated by the electrophoretic data.

This may be due to the small sample of enzymes, as a consequence of the abundant resin in the winter buds of the *Tacamahaca* species. Indeed, other enzyme systems such as

6-Phosphoglucose Isomerase and Glucose 6-Phospate Dehydrogenase have proved to be useful in the distinction of closely related *Leuce* taxa (Aleksiuk and Barnes, unpublished). These two enzymes and several others could not be used here because their activity was suppressed by the acetone used in cleaning the buds.
Another possible explanation may be the poor differentiations of the taxa studied or introgression.
If not efficient at the species level, this technique however did allow the identification of pure *P. balsamifera* locations from locations with species mixtures.
Also, the extremely abundant variation in the genotypes between the individuals (figure 24 and appendix 4), within- and between-locations, allows one to conclude to the existance of a high polyclonality in the general sample.

5 - Intra-individual variability of the allozyme pattern

To check the within-individual variability of the allozyme pattern, samples were collected on different parts of the same tree. The selected tree was located near Spruce Grove (25km West of Edmonton). Samples were collected on the different sides of the tree (north, south, east and west) and in three different part of the crown (upper crown, middle branches and basal branches). Two more samples were

collected from suckers near the base of the trunk. None of the 14 samples showed any variation in the allozyme pattern, including the sucker samples.

The variability of the allozyme pattern throughout the year was also examined. On a tree located in Edmonton (Saskatchewan River Valley), bud samples were collected every week from September 3rd 1987, to April 28th 1988. None of these samples showed any variation in the allozyme pattern, even the early samples which were especially resinous and the last samples that were partially flushed.

To check the stability of the allozyme pattern from year to year, collections were made twice at Kananaskis 4, first in the winter of 1987 with the rest of the general sample, and then during the winter of 1988. The allozyme patterns of the 20 trees sampled in this location showed no variation between the two collections of 1987 and 1988.

VIII - General discussion and conclusions

1 - Method comparison

A constant in *Populus* literature is the report of high within-individual variability in morphological characters. Some of this variability can be controlled by systematically using "early leaves" collected on one- or two-year-old shoots in a particular part of the crown. During this controlling some of the information may be lost. By eliminating certain categories of leaf material because of their high within-individual variability, one may also eliminate some of the potentially useful between-individual or interspecific variation.

In this study, by using morphological characters, it was possible to distinguish *P. angustifolia* from *P. balsamifera* as well as their putative hybrids by correspondence analysis. The two *P. balsamifera* subspecies could not be separated by this method. However, trends towards one or the other subspecies can be seen using correspondance analysis (figures 13 and 14). Between these two trends, there is a continuous distribution of the individuals denoting an almost continuous variation of the leaf shape in *P. balsamifera* in southwestern Alberta.

Within-individual variability is not the only limiting factor of the morphological approach. Great

between-individual variability in a species can also cause problems when dealing with closely related taxa. Barnes (1978), reported great variation in morphological characters in *P. tremuloides*, even within a relatively localized sample; he suggested that this variation could mask some of the gene flow occurring between *P. tremuloides* and *P. grandidentata*. He also suggested the use of other techniques such as chemical traits and pollen viability to help elucidate relationships.

The use of floral morphological characters and pollen viability was limited by the extent of flowering in the population studied. In 1989, only a third of the general sample could be investigated. Several more years of field collection would be necessary to cover the entire population, of which a part might never flower (Brayshaw 1965a, 1965b).

With a limited number of individuals tested, these two techniques revealed some insights on *P. balsamifera* and its two subspecies.

The female catkins examined consisted of a mix of two- and three-carpellate flowers. An important proportion (60%) of the male individuals had a number of stamens ranging from 20 to 40. These observations are characteristic of *P. balsamifera* individuals intermediate between subsp. *balsamifera* and subsp. *trichocarpa* (Brayshaw 1965a, Vierek and Foote 1970).

Pollen viability results suggest the same conclusion. The relatively low pollen viability in the individuals tested suggests the occurrence of hybridization between *P. angustifolia* and *P. balsamifera* and possibly between the two *P. balsamifera* subspecies.

Of the three techniques investigated in this study, electrophoresis gave the least satisfying separation of the three taxa studied. Electrophoresis was seriously limited by the abundant resin present in the winter buds. Root material could not be utilized because of intra-individual variability. As a result, only nine enzyme systems could be investigated. If not efficient at separating taxa, the method permitted the identification of locations having only *P. balsamifera* (both supspecies), from locations having only *P. angustifolia*, and also from locations having both species plus putative hybrids.

Populus is well known for its ability to reproduce vegetatively by suckering (Dickman and Stuart 1983) especially in section *Leuce* (Barnes 1969). In this study, electrophoresis showed a high between-individual variability in most locations. This variability confirms high polyclonality in American balsam poplar populations reported by Comtois et al. (1986).

Flavonoid chemistry, does not present the limitations of the other techniques investigated. Leaf material can be collected at any time of the growing season and stored indefinitely. No intra-individual variability was found. Although flavonoid chemistry permitted the separation of the three taxa, differences between *P. angustifolia* and *P.balsamifera* were much more distinct than those between subspecies of *P. balsamifera*. *P. angustifolia* differs from *P. balsamifera* mainly by the absence of flavones. The two subspecies of *P. balsamifera* differ in their glycosylation pattern, subsp. *balsamifera* showing a less complex pattern than subsp. *trichocarpa*. About a third of the individuals of the general sample exhibit pure profiles (figure 19). The rest show several levels of intermediacy between pure taxa profiles (figure 21). Three different types of intermediate profiles between *P. angustifolia* and *P. balsamifera* were distinguished. Between the two *P. balsamifera* subspecies, three intermediate profiles exist, one tends more towards subsp. *balsamifera* and one towards subsp. *trichocarpa*. Also, in many *P. balsamifera* samples, *P. angustifolia* influence is revealed by the suppression of one or two flavone compounds from their profiles. No sign of this influence could be observed when examining leaf morphology. Individuals exhibiting intermediate flavonoid patterns should be considered as putative hybrids. Also, the various levels of intermediacy suggest the possible occurrence of back-crossing and subsequent gene flow in southwestern

Albertan populations.

In southwestern Alberta, *P. balsamifera* subsp. *balsamifera* is more strongly representated than subsp. *trichocarpa*. This is revealed by a greater number of individuals having flavonoid profiles typical of this subspecies or tending towards it. Furthermore, floral morphology shows that male individuals with smaller numbers of stamens (12 to 20) characteristic of subsp. *balsamifera*, are more numerous in such populations.

Brayshaw (1965a) in southwestern Alberta and Viereck and Foote (1970) in Alaska reported that morphological differences were not sufficient to treat the two subspecies of *P. balsamifera* as two distinct species. In this study, it was found that their flavonoid patterns possess great similarities in aglycone composition, but differ in glycosylations.

Subsp. *balsamifera* and subsp. *trichocarpa* correspond to the definition of subspecies by Stebbins (1950). They have morphological and phytochemical characters in common, such as leaf shape and aglycone pattern. They differ in several other characteristics such as flower and fruit characters and glycosylation pattern. In southwestern Alberta, there are many intermediate forms which may indicate a lack of efficient breeding isolation mechanisms between the two taxa, to prevent gene flow between populations. These taxa should not be considered separate species but at most

subspecies as Brayshaw (1965a) suggested in southwestern Alberta. Hultèn (1968a, 1968b) extended the subspecies status to the whole range of the two taxa in North America. A general survey of the flavonoid pattern of balsam poplars in North America would be useful to confirm this extended subspecific status.

2 - Update of Brayshaw's distribution maps

During the summer of 1985, the locations where Brayshaw collected in 1959 and 1963 (1965b) were visited; most of them could be found. Brayshaw did not give any indications as to size, age and health of the poplar stands he visited and comparison with his actual observations is not possible. However, it was noticed that in several locations *P. angustifolia* had disappeared or was disappearing. Also, several putative hybrid combinations determined by the author could not be found. For example, in Taber, the population consisted only of *P. deltoides*, *P. angustifolia* and *P. balsamifera* could not be found. This might indicate a diminution in size of the locations and/or the non-regeneration of certain species. In several areas, such as Writing-on-Stone Park and St Mary Reservoir, the native poplars were replaced by cultivated varieties.

Brayshaw's distribution maps (1965b) generally are still accurate. One exception is the presence of *P. angustifolia* along Willow Creek (cf. figure 6 and 7) which the author apparently never visited. In several other locations such as Whyndam-Carseland and Chain Lakes, *P. angustifolia* was found. In these two locations *P. angustifolia* individuals were overgrown by the rest of the population and dying, but several putative hybrids were present.

The *P. balsamifera* distribution map has not changed much since 1963. However, a difference between populations located in the prairie and those located in the foothills was noticed. Foothills populations are thriving with abundant regeneration indicated by the presence of numerous young individuals, even in high altitude locations such as Waterton and Kananaskis. Contrary to this, prairie populations, especially in the southern part of the area, showed poor regeneration with many dead or dying trees.

The southwestern Alberta prairie has been modified by progressive damming of the river systems. Two important dams on the St Mary River and Waterton River were completed in 1951 and 1964 respectively. A third is under construction on the Oldman River. Rood (1988) and Rood and Mahoney (1989) studied the evolution of the poplar populations along the

two dammed rivers, using two sets of air photos taken in 1961 and 1981. During this 20-year period, the poplar populations were found to have declined by 48% on St Mary River and by 23% on Waterton River, downstream from the dams. On the other hand, poplar populations situated upstream from the dams showed only a slight decline. The authors also studied the Belly River which is not dammed and runs almost parallel to the two other rivers. Very little variation was found on this river.

The authors attributed this decline to the shutting off of water during the dry summer months when the rivers downstream of the dams are reduced to no more than a narrow creek. Spring flooding still occurs and permits poplar seedling establishment, but the river level during the summer months is too low for the survival of these seedlings. With the construction of a new dam on the Oldman River there is concern about the future of the river gallery forests in southwestern Alberta. Several populations used in this study might disappear in the next years. This process could also accelerate the elimination of *P. angustifolia* from Alberta where it is scarce in its pure form (Brayshaw 1965b), because of its intolerance to competition and obligatory riparian habitat.

ASAKAWA Y. & WOLLENWEBER E. - 1976
A novel phenolic acid derivative from buds of *Populus lasiocarpa*. Phytochemistry 15 : 811-812.

AVISE J.C. - 1974
Systematic value of electrophoretic data. Systematic Zoology 23 : 465-481.

BARNES B.V. - 1969
Natural variation and delineation of clones of *Populus tremuloides* and *P. grandidentata* in northern lower Michigan. Silvae Genetica 18 : 130-142.

BARNES B.V. - 1975
Phenotypic variation of trembling aspen in western North America. Forest Science 21 : 319-329.

BARNES B.V. - 1978
Pollen abortion in *Betula* and *Populus* (section *Leuce*). Michigan Botanist 17 : 167-172.

BATE-SMITH E.C. & LERNER N.H. - 1954
Leuco-anthocyanins. 2- Systematic distribution of leuco-anthocyanins in leaves. Biochemical Journal 58 : 126.

BATE-SMITH E.C. - 1962
Phenolic constituents of plants and their taxonomic significance. I - Dicotyledones. Journal of the Linnean Society (Botany) 58 : 95-173.

BATE-SMITH E.C. - 1968
The phenolic constituents of plants and their taxonomic significance. II - Monocotyledons. Journal of the Linnean Society (Botany) 60 : 325-356.

BENZECRI J.P. - 1969
Statistical analysis as a tool to make patterns emerge from data. *in* S. Watanabe (ed.) : Methodologies of pattern recognition. Academic Press, New York, pp.35-74.

BERGMAN F. - 1987
Characterization of multiclonal aspen cultivars using isozyme electrophoresis. Forestry Ecology and management 22 : 167-172.

BOLD H.C. - 1967
Morphology of plants, Harper & Row publishers, New York, 3rd edition, 668p.

BRAYSHAW T.C. - 1965a
The status of the Black Cottonwood (*Populus trichocarpa* Torrey and Gray). Canadian Field Naturalist 79 : 91-95.

BRAYSHAW T.C. - 1965b
Native poplars of southern Alberta and their hybrids. Department of Forestry Publication n°1109, 40p.

BREWER G.J. - 1970
An introduction to isozyme techniques, Academic Press, New York and London, 186p.

BRITTON N. - 1923
Illustrated flora of the northern United States, Canada and the British possessions, Scribner, New York, 3 Vol..

BUTH D.G. - 1984
The application of electrophoretic data in systematic studies. Annual Review of Ecology and systematics 15 : 501-522.

CASTILLO T. & PADRO A. - 1987
Electrophoretic characterization of the Euramerican poplar clones "I-214" and "Campeador". Silvae Genetica 36 : 250-251.

CEULEMANS R., IMPENS I. & STEENACKERS V. - 1987
Variations in photosynthetic, anatomical and enzymatic leaf traits and correlations with growth in recently selected *Populus* hybrids. Canadian Journal of Forest Research 17 : 273-283.

CHADENSON M., HAUTEVILLE M., CHOPIN J., WOLLENWEBER E., TISSUT M. & EGGER K. - 1971
Identification d'un nouveau produit naturel, le diOH 2,6 métoxy-4 dibenzoylméthane, extrait des bourgeons du peuplier *Populus nigra*. Comptes Rendus Hebdomadaires de l'Académie des Sciences Série D 273(25) : 2658-2660.

CHARRIERE-LADREIX Y. - 1973
Flavonoid secretion of *Populus nigra* var. *italica* buds. Kinetics of secretory phenomena, ultra-structure and glandular development. Journal de Microscopie 17 : 299-316.

CHARRIERE-LADREIX Y. - 1976
Répartition intra-cellulaire du sécrétat flavonique de *Populus nigra* L.. Planta 129(2) : 167-174.

CHELIAK W.M. & PITEL J.A. - 1984a
Electrophoresis identification of clones in trembling aspen. Canadian Journal of Forest Research 14 : 740-743.

CHELIAK W.M. & PITEL J.A. - 1984b
Techniques for starch gel electrophoresis of enzymes from forest tree species. Petawawa National Forestry Institute Information Report PI-X-42, Canadian Forestry Service, Agriculture Canada, 49p.

COLLIER M.D., MURRAY D.R. - 1977
Leucyl-betanaphthylamidase activities in developing seeds and seedlings of *Pisum sativum* L.. Australian Journal of Plant Physiology 4 : 571-582.

COMTOIS P., SIMON J.P. & PAYETTE S. - 1986
Clonal constitution and sex ratio in northern populations of balsam poplar *Populus balsamifera*. Holartic Ecology 9 : 251-260.

CONKLE M.T., HODGSKISS P.D., NUNNALLY L.B. & HUNTER S.B. - 1982

Starch gel electrophoresis of conifer seeds : a laboratory manual. U.S.D.A. Forest Service. Pacific Southwest Forest and Range Experiment Station, General Technical Report PSW-64, 18p.

CRAWFORD D.J. - 1974
A morphological and chemical study of *Populus acuminata*. Brittonia 26 : 74-89.

CRITCHFIELD W.D. - 1960
Leaf dimorphism in *Populus trichocarpa*. American Journal of Botany 47 : 699-711.

DE CANDOLLE A.P. - 1816
Essai sur les propriétés médicales des plantes, comparées avec leurs formes extérieures et leur classification naturelle, Paris, 2ème édition.

DE JUSSIEU B. - 1741
Histoire d'une plante connue sous le nom de *Pilularia*. Mémoire de l'Académie Royale des Sciences, Paris 1739 : 240-256.

DICKMAN D.I. & STUART K.W. - 1983
The culture of poplars in eastern North America. Michigan State University Press, Dansville, Michigan, 167p.

ECKENWALDER J.E. - 1977
Systematics of *Populus* L. (*Salicaceae*) in southwestern North America with special reference to section *Aigeiros* Duby. Ph.D. thesis, University of California, Berkeley.

EGGER K. & TISSUT M. - 1968
Sur la présence de galangine, pinocembrine et izalpine dans les bourgeons de *Populus nigra* var. *italica*. Comptes Rendus Hebdomadaires de l'Académie des Sciences, série D 267(16) : 1329-1332.

EGGER K., TISSUT M. & WOLLENWEBER E. - 1969
3-O Methylather der Kaempferols und der galangins in Knospenol von *Populus nigra*. Phytochemistry 8 :

2425-2426.

F.A.O. - 1980
Poplars and willows in wood production and land use. F.A.O. Forestry series n°10, 328p.

FERET P.P. - 1971
Isoenzyme variation in *Picea glauca* (Moench) Voss seedlings. Silvae Genetica 20 : 46-50.

FINESCHI S., MALVOTI M.E. & VILLANI F. - 1986
Differentiation of poplar clones through isoenzyme patterns. Proceedings of 18th IUFRO World Congress, Ljubljana, Yugoslavia, Div. 2, Forest plants and forest protection, Vol. 2, pp.832.

FOWELLS H.A. - 1965
Silvics of forest trees of the United States. U.S. Department of Agriculture Handbook 271, 762p.

FUTUYMA D.J. - 1986
Evolutionnary Biology, 2nd edition, Sinauer Associates Inc., Sunderland, Massachussetts, 600p.

GRANT V. - 1971
Plant speciation. Columbia University Press, New York and London, 435p.

GREENACRE M.J. & DEGOS L. - 1977
Correspondence analysis of HLA gene frequency data from 124 population samples. American Journal of Human Genetics 29 : 60-75.

GREENACRE M.J. - 1984
Theory and applications of correspondence analysis. London Academic Press Inc., 364p.

GREENAWAY W., SCAYSBROOK T. & WHATLEY F.R. - 1988
Phenolic analysis of bud exudate of *Populus lasiocarpa* by GC/MS. Phytochemistry 27(11) : 3513-3515.

GUZINA V. - 1974
Isoenzymes of peroxidases in genetic studies of

poplars. Genetika 6 : 63-68.

HARBORNE J.B. - (1966)
Comparative biochemistry of flavonoids. I- Distribution of chalcone and aurone pigments in plants. Phytochemistry 5 : 111-115.

HARBORNE J.B. - (1967)
Comparative biochemistry of the flavonoids. Academic Press, London.

HARBORNE J.B. - 1973
Flavonoids. *in* Miller L.P. : Phytochemistry, Van Nostrand Reinhold Company, New York, Vol.2, pp. 345-380.

HARBORNE J.B., MABRY T.J. & MABRY H. - 1975
The flavonoids. Chapman and Hall, London, 1204p.

HARRIS H. - 1969
Enzyme and protein polymorphisms in human populations. British Medical Bulletin 25 : 5-13.

HAUTEVILLE M., CHADENSON M. & CHOPIN J. - 1973
Un nouveau type de flavonoides naturels : les dihydroxy-2,5 flavanones. I-Synthèse du "dihydroxy 2,6 metoxy-4 dibenzoylméthane" des bourgeons de *Populus nigra* et mise en évidence de sa structure cyclique. Bulletin de la Société Chimique de France 5 : 1781-1784.

HEGNAUER R. - 1973
Chemotaxonomie der pflanzen, *Salicaceae*, Vol. 6, ch.2.2.3, Birkhauser verlag Basel und Stuttgart pp.241-258.

HILL B.M. - 1973
Reciprocal averaging : an eigenvector method of ordination. Journal of Ecology 61 : 237-249.

HILL M.O. - 1974
Correspondence analysis : a neglected multivariate method. Applied statistics 23(3) : 340-354.

HILL M.O. - 1982
Correspondence analysis. *In* S. Kotz & N.L. Johnson (eds) : Encyclopedia of statistical Sciences. John Wiley and sons Inc., New York, Vol.2, pp.204-210.

HIRSHFIELD H.O. - 1935
A connection between correlation and contingency. Proceedings of the Cambridge Philosophical Society 31(Oct.) : 520-524.

HITCHCOCK C.L., CRONQUIST A, OWNBEY M. & THOMPSON J.W. - 1964
Vascular plants of the Pacific Northwest. Part 2 : *Salicaceae* to *Saxifragaceae*. University of Washington Press, Seattle & London, 597p.

HOFFMAN D.L. & FRANKE G.R. - 1986
Correspondence analysis : graphical representation of categorical data in marketing research. Journal of Marketing Research 23(August) : 213-227.

HOSIE R.C. - 1961
Native trees of Canada. Canadian Forestry Service, Department of Fisheries and Forestry, 291p.

HU C., CROVELLO T.J. & SOKAL R.R. - 1985
The numerical taxonomy of some species of *Populus* based only on vegetative characters. Taxon 34 : 197-206.

HULTEN E. - 1968a
Comments on the flora of Alaska and Yukon. Arkiv for Botanik 7(1) : 1-147.

HULTEN E. - 1968b
Flora of Alaska and neighbouring countries. A manual of the vascular plants. Stanford University Press, Stanford, California, 1008p.

JAMES E. - 1823
Long's Expedition. Account of an expedition from Pittsburg to the Rocky Mountains, performed in the years 1819-1820, Vol I, pp.497.

JONES A.J. & SEIGLER D.S. - 1974
Populational observations and flavonoid data in support of hybrid status for *Populus acuminata*. American Journal of Botany 61(5suppl.) : 45.

KIM C.S. & CHUNG S.S. - 1974
Variation in the patterns of isoperoxidase in *Populus alba*, *P. glandulosa* and *P. x euramericana*. Research Report, Institute of Forest Genetics 11 : 53-59.

LEBART L, MORINEAU A. & WARWICK K.M. - 1984
Multivariate descriptive statistical analysis : correspondence analysis and related techniques for large matrices. Wiley, New York, 231p.

LEBRETON P. - 1962
Contribution à l'étude des flavonoides chez *Humulus lupulus* L. et autres urticales. Thèse Sciences Naturelles, Lyon.

LESTER D.T. - 1963
Variation in sex expression in *Populus tremuloides* Michx.. Silvae Genetica 12 : 141-151.

LEWONTIN R.C. & HUBBY J.L. - 1966
A molecular approach to the study of heterozygosity in natural populations. II-Amount of variation and degree of heterozygosity in natural populations of *Drosophila pseudoobscura*. Genetics 54 : 595.

LINNEAUS C. VON - 1753
Species plantarum, pl. 1034.

LITTLE E.L. - 1971
Atlas of United States Trees. Vol I, Conifers and important hardwoods. U.S. Forest Service Miscellaneous Publication N°1146.

LITTLE E.L. - 1976
Atlas of United States Trees. Vol 3, Minor western hardwoods. U.S. Forest Service Miscellaneous Publication N°1314.

MABRY T.J., MARKHAM K.R. & THOMAS M.B. - 1970
The systematic identification of flavonoids. Springer Verlag, Berlin, Heidelberg, 354p.

MARKHAM K.R. - 1982
Techniques of flavonoid identification. Academic Press, London, 113p.

MOHRDIEK O. - 1983
Discussion : future possibilities for poplar breeding. Canadian Journal of Forest Research 13 : 465-471.

MOSS E.H. - 1932
The vegetation of Alberta. IV-The poplar association and related vegetation of central Alberta. Journal of Ecology 20 : 380-415.

MUHLE LARSEN C. - 1970
Recent advances in poplar breeding. International Review of Forestry Research 3 : 1-67.

PACKER J.G. - 1983
Flora of Alberta, 2nd ed. (revised from E.H. Moss). University of Toronto Press, 687p.

PAULEY S.S. - 1949
Forest tree genetics research : *Populus* L.. Economic Botany 3 : 299-330.

PEARL I.A. & DARLING S.F. - 1971
Phenolic extractives of the leaves of *Populus balsamifera* and of *P. trichocarpa*. Phytochemistry 10(11) : 2844-2847.

RADFORD A.E., DICKISON W.C., MASSEY J.R. & BELL C.R. - 1974
Vascular plant systematics. Harper & Row, New York, 891p.

RAJORA O.P. - 1986
Studies on genetics and relationships of *Populus deltoides*, *P. nigra* and *P. maximowiczii* using isozymes, pollen competition and leaf morphology. Ph.D. thesis, University of Toronto, Toronto.

RAJORA O.P. - 1988
Allozymes as aids for identification and differentiation of some *Populus maximowiczii* Henry clonal varieties. Biochemical Systematics and Ecology 16(7/8) : 635-640.

RAJORA O.P., ZSUFFA L. - 1989
Multilocus genetic structure, characterization, and relationships of *Populus x canadensis* cultivars. Genome 32(1) : 99-108.

REHDER A. - 1951
Manual of cultivated trees and shrubs hardy in North America. 2nd edition, Mac Millan, New York, 996p.

RIBEREAU-GAYON P - 1968
Les composés phénoliques des végétaux, Dunod, Paris.

RIDGWAY G.J., SHERBURNE S.W. & LEWIS R.D. - 1970
Polymorphisms in the esterases of Atlantic herring. Transactions American Fisheries Society 99 : 147-151.

ROOD S.B. - 1985
Clarifying the status of *Populus* in southern Alberta. University of Lethbridge Research Note, 13p.

ROOD S.B., CAMPBELL J.S. & DESPINS T. - 1986
Natural poplar hybrids from southern Alberta. I-Continuous variation for foliar characteristics. Canadian Journal of Botany 64 : 1382-1388.

ROOD S.B. & MAHONEY J.M. - 1989
The collapse of river valley forests downstream from dams in the Rocky Mountains foothills. Proceedings, Canadian Society of Environmental Biology (in press).

S.A.S. - 1985
S.A.S. User's guide : Basics. Version 5 edition. Cary, NC : S.A.S. Institute Inc., 1290p.

S.A.S. - 1985
S.A.S. User's guide : Statistics. Version 5 edition. Cary, NC : S.A.S. Institute Inc., 956p.

SCHAAL B.A. & ANDERSON W.W. - 1974
An outline of techniques for starch gel electrophoresis of enzymes from the American oyster *Crassostrea viriginca* Gmelin. Georgia Marine Science Center, Technical Report Series n°74-3, 17p.

SHAW C.R. & PRASAD R. - 1970
Starch gel electrophoresis - A compilation of recipes. Biochemical Genetics 4 : 297-320.

SICILIANO M.J. & SHAW C.R. - 1976
Separation and visualization of enzymes on gels. *in* I. Smith : Chromatographic and electrophoretic techniques, Vol 2. Zone electrophoresis. Heinemann, London, pp 185-289.

SMITH P.M. - 1976
The chemotaxonomy of plants. Contemporary Biology, Arnold Edward, London.

SOKAL R.R., CROVELLO T.J. & UNNASCH R.S. - 1986
Geographic variation of vegetative characters of *Populus deltoides*. Systematic Botany 11(3) : 419-432.

SOSA F. & PERCHERON F. - 1970
Le Kaempférol 3-O rhamno-diglucoside, nouvel hétéroside du pollen du *Populus yunnanensis*. Phytochemistry 9(2) 441-446.

STEBBINS G.L. - 1950
Variation and evolution in plants. Columbia University Press, New York, 643p.

STETTLER R.F. - 1965
Sex expression in *Populus trichocarpa* (T.&G. ex Hook.) and its hybrids. Proceedings Western Forest Genetics Association, Olympia, Washington, 6-7(december) : 45-53.

STETTLER R.F. - 1971
Variation in sex expression of Black Cottonwood and related hybrids. Silvae Genetica 20(1-2) : 42-46.

STOUT A.B. & SCHREINER E.J. - 1933
Results of a project in hybridizing poplars. Journal of Heredity 24 : 216-229.

TAMAS M., MARINESCU I. & IONESCU F. - 1979
Flavonoidele din mugurii de plop. Studii si Cercetari de Biochimie 22(2) : 207-213.

TEIL H. - 1975
Correspondance factor analysis : an outline of its method. Mathematical Geology 7(1) : 3-12.

THIEME H. & RICHTER R. - 1966
Isolation of a new glycoside from *Populus tremula*. Pharmazie 21 : 251.

THIEME H. - 1969
Die flavonoide der Salicaceen. III-Flavonoide der Blatter von *Populus candicans* Ait.. Pharmazie 24 : 783-784.

TORREY J. & GRAY A. - 1843
Flora of North America. Wiley and Putnam, New York, 2 Vol..

VANENDE A. - 1984
La variation polyphénolique des exsudats de bourgeons chez *Populus trichocarpa* Torrey & Gray. Thèse de 3ème Cycle. Université de Paris-Sud Orsay, Orsay.

VIEREK L.A. & FOOTE J.M. - 1970
The status of *Populus balsamifera* and *P. trichocarpa* in Alaska. Canadian Field Naturalist 84(2) : 169-173.

VILAR M., GAGET M. & DUMAS C. - 1988
Micro-isoelectric focusing of proteins from single stigmas of *Populus*. Canadian Journal of Forest

Research 18 : 1261-1264.

WANG M., HUANG M. & al. - 1982
Isoenzyme analysis in clone identification of *Aigeiros* poplars. Journal, Nanjing Technological College of Forest products 1 : 105-111.

WANG M., HUANG M. & al. - 1983
A genetic analysis of peroxydase isoenzymes in poplars. Journal, Nanjing Technological College of Forest Products 3 : 137-140.

WEBER J.C. - 1980
Isoenzyme variation among ten populations of *Populus trichocarpa* (Torr. & Gray) in the Pacific Northwest. Master of Science, University of Washington, Seattle, Washington.

WEBER J.C. & STETTLER R.F. - 1981
Isoenzyme variation among ten populations of *Populus trichocarpa* Torr. & Gray in the Pacific Northwest. Silvae Genetica 30(2-3) : 82-87.

WOLLENWEBER E. - 1974
Methylated flavonols from buds of several species of *Populus*. Phytochemistry 13 : 760-761.

WOLLENWEBER E. - 1975
Flavonoidmuster als systematisches merkmal in der gattung *Populus*. Biochemical Systematics and Ecology 3(1) : 35-45.

WOLLENWEBER E. & EGGER K. - 1971
Die lipophilen flavonoide der knoojenols von *Populus nigra*. Phytochemistry 10(1) : 225-226.

WOLLENWEBER E. & WEBER W. - 1973
Die farbgebenden chalkone des pappelols. Zeitschrift fur Pflanzenphysiologie 69(2) : 125-128.

YEH F.C.H. & O'MALLEY D. - 1980
Enzyme variations in natural populations of Douglas fir *Pseudotsuga menziesii* (Mirb.) Franco,

from British Columbia. 1- Genetic variation patterns in coastal populations. Silvae Genetica 29 : 83-92.

ZSUFFA L. - 1975
A summary review of interspecific breeding in the genus *Populus* L.. Proceedings, Meeting of the Canadian Tree Improvement Association 14(2) : 107-123.

Appendix 1 - Correspondence analysis program.

```
FILE: CORRESP  DOC      *                    VM/SP CONVERSATIONAL MONITOR SYSTEM
CORRESP is a SAS MACRO to perform correspondence analysis (CA).  CA can
be thought of as a principal components analysis of a contingency table
in which both the rows and columns are analyzed simultaneously.  The
resulting component scores . .

Input data can be in any of several forms:
 1) contingency table
    with rows and columns being the levels of one or more variables
 2) cases by variables matrix
    a) cases are subjects
    with rows being subjects and columns being levels of variables
    b) cases are groups of subjects

To run the macro, add the "MACRO" and "MAUTOSOURCE" options to the
SAS command, e.g.
 SAS commandfilename (MACRO MAUTOSOURCE


Invoke the macro in a SAS commandfile with the following statement:
 %corresp()                    to not specify any of the macro's options
or
 %corresp(macro options)       to specify some of the macro's options


The following options (with their defaults) are available and
should be separated by commas:

 DATA=_LAST_             /* which data set to analyze
                            /* if none, the last one created is used
 VAR=_NUMERIC_            /* which variables to analyze
                            /* if none, all numeric variables are used
 ID=                      /* the character variable to use as a row id:
                            /* if none, one is create "OBSnnn"
 N=                       /* maximum number of components to retain
                            if none, all components retained
 CONTRIB=YES              /* absolute and relative contributions are
                            YES: are calculated and printed
                            NO: are not calculated and printed
 CORRECT=NO               /* Carroll, Green, Schaffer correction:
                            allows comparison of distance between
                               row score and column score
                            YES: is used
                            NO: is not used
 BURT=NO                  /* use BURT matrix: premultipy data by transpose:
                            used to analyze large cases by variables
                            matrix (row scores are lost, but see
                            ROWRECON option)
                            YES: is used
                            NO: is not used
 ROWRECON=NO              /* reconstruct row scores lost from use of
                            BURT matrix (only to be used with BURT=YES)
                            YES: row scores reconstructed
                            NO: row scores not reconstructed


Data sets created

 EIGEN     eigenvalues(inertia) and proportion of total inertia
 CCOORD    component scores for columns
 RCOORD    component scores for rows
           NOT calculated for (BURT=YES and ROWRECON=NO)
 CABSCON   absolute contributions for columns
           NOT calculated for CONTRIB=NO
 RABSCON   absolute contributions for rows
           NOT calculated for CONTRIB=NO
 CRELCON   relative contributions for columns
           NOT calculated for CONTRIB=NO
 RRELCON   relative contributions for rows
           NOT calculated for CONTRIB=NO
```

```
FILE: CORRESP  SAS          *                VM/SP CONVERSATIONAL MONITOR SYSTEM
OPTIONS NOSOURCE;
%MACRO     CORRESP                 /* CORRESPONDENCE ANALYSIS              */
                                   /*-------------------------------------*/
                                   /* DEFINE PARAMETER DEFAULTS            */
                                   /*-------------------------------------*/
          (DATA = _LAST_,          /* DATA SET OF VARIABLES                */
           VAR = _NUMERIC_,        /* LIST OF VARIABLES FOR ANALYSIS       */
           ID = ,                  /* OBSERVATION ID CHARACTER VARIABLE    */
           N = ,                   /* MAXIMUM NUMBER OF COMPONENTS TO RETAIN */
           CORRECT = NO,           /* YES,NO: perform CGS correction       */
           CONTRIB = YES,          /* YES,NO: print abs & rel contributions */
           BURT = NO,              /* YES,NO: use BURT matrix for data     */
           ROWRECON = NO);         /* YES,NO: reconstruct row scores when
                                      using BURT matrix                    */
%LET CORRECT=%UPCASE(&CORRECT);
%LET CONTRIB=%UPCASE(&CONTRIB);
%LET BURT=%UPCASE(&BURT);
%LET ROWRECON=%UPCASE(&ROWRECON);

%LET OPTERR=NO;

%IF &BURT NE NO AND &BURT NE YES %THEN %DO;
%PUT ERROR: Illegal value for BURT parameter (&BURT);
%LET OPTERR=YES;
%END;

%IF &CORRECT NE NO AND &CORRECT NE YES %THEN %DO;
%LET OPTERR=YES;
%PUT ERROR: Illegal value for CORRECT parameter (&CORRECT);
%END;

%IF &CONTRIB NE NO AND &CONTRIB NE YES %THEN %DO;
%LET OPTERR=YES;
%PUT ERROR: Illegal value for CONTRIB parameter (&CONTRIB);
%END;

%IF &ROWRECON NE NO AND &ROWRECON NE YES %THEN %DO;
%LET OPTERR=YES;
%PUT ERROR: Illegal value for ROWRECON parameter (&ROWRECON);
%END;

   /* terminate macro if option errors */
%IF OPTERR=YES %THEN %GOTO OUT;

*-------------------------------------------------------------------------*
|    CREATE _NOMISS_ THROUGH CASEWISE DELETION OF MISSING DATA.           |
|-------------------------------------------------------------------------|
|----CREATE--COMP--DATA-SET-WITH-"COMP1"-"COMP2"-ETC----------------------;|

DATA _NOMISS_(DROP = I _CNAME_)
    _COMP_(KEEP = _CNAME_);
   SET  &DATA;
   LENGTH _CNAME_ $ 8;
   IF _N_ = 1 THEN DO;
      ARRAY _Y_(*) &VAR;
      DO I=1 TO DIM(_Y_);
         _CNAME_ = COMPRESS("COMP"||PUT(I,4.));
         OUTPUT _COMP_;
         END;
```

```
      END;
    ELSE I=1;        /* keeps I from being missing for &VAR=_numeric_ */
    IF NMISS(OF &VAR) = 0;
%IF &ID=    %THEN %DO;
    _RNAME = COMPRESS("OBS"||PUT(_N_,4.));
%LET ID=_RNAME;
%END;
    OUTPUT _NOMISS_;
    KEEP &VAR &ID _CNAME_;
RUN;

PROC MATRIX FUZZ;
*************************************;
NOTE skip=2 'OPTIONS IN EFFECT';
%IF &N= %THEN %DO;
note skip=2 "(N=MAX): Maximum number of requested components";
%END;
%ELSE %DO;
note skip=2 "(N=&N): Maximum number of requested components";
%END;

%IF &BURT=NO %THEN %DO;
NOTE skip=2
'(BURT=NO): data matrix used as is (NOT pre-multiplied by its transpose)';
%END;
%ELSE %IF &BURT=YES %THEN %DO;
NOTE skip=2
'(BURT=YES): data matrix is pre-multiplied by its transpose          ';
NOTE '(useful with large ''subjects by variables'' data matrix)';
%END;

%IF &CORRECT=NO %THEN %DO;
NOTE skip=2
'(CORRECT=NO): Carroll, Green, Schaffer correction NOT applied       ';
%END;
%ELSE %IF &CORRECT=YES %THEN %DO;
NOTE skip=2
'(CORRECT=YES): Carroll, Green, Schaffer correction applied';
NOTE '(allows comparison of distances between row point and column point)';
%END;

%IF &CONTRIB=NO %THEN %DO;
NOTE SKIP=2
'(CONTRIB=NO): Absolute and relative contributions NOT printed       ';
%END;
%ELSE %IF &CONTRIB=YES %THEN %DO;
NOTE SKIP=2
'(CONTRIB=YES): Absolute and relative contributions are printed      ';
%END;
*************************************;
FETCH X DATA=_NOMISS_ ROWNAME=&ID COLNAME=CNAME;
RNAME=&ID; free &ID;
%IF &BURT=YES %THEN %DO;
 X=%STR(X%')*X; RNAME=%STR(CNAME%');
 %END;

%if &N = MAX OR &N =  %THEN %DO;
MAXC=MIN(NROW(X)-1||NCOL(X)-1);
%END;
%ELSE %DO;
```

```
MAXC=MIN(NROW(X)-1||NCOL(X)-1||&N);
%END;

P=X#/SUM(X);
FREE X;
R=P(,+);
C=P(+,);
/* COMMENTED STATEMENT IS HOFFMAN ORIGINAL.
  IT HAS BEEN REPLACED BY FOLLOWING STATEMENT
Y=INV(DIAG(SQRT(R))) * (P-R*C) * INV(DIAG(SQRT(C)));
*/
Y=DIAG(1#/SQRT(R)) * P * DIAG(1#/SQRT(C));
FREE P;
SVD U S V Y;
FREE Y;
%IF &BURT=YES %THEN %DO;
 S=SQRT(S);
%END;
RNK=RANK(-S);
TEMP=U; U(,RNK)=TEMP;
TEMP=S; S(RNK,)=TEMP;
TEMP=V; V(,RNK)=TEMP;
FREE TEMP RNK;
/* COMMENTED STATEMENTS ARE HOFFMAN ORIGINAL.
  THEY HAVE BEEN REPLACED BY FOLLOWING STATEMENTS
 POS=SUM(S>.000001); S=S(1:POS,); U=U(,1:POS); V=V(,1:POS);
*/
POS=SUM(S>.000001);
S=S(2:POS,); U=U(,2:POS); V=V(,2:POS);

FETCH COMPNAME DATA=_COMP_ TYPE=CHAR;

maxc=min(nrow(S)||maxc);        /* max # components to retain */

EIGEN=S##2;
PROP=EIGEN#/SUM(EIGEN); PROP=PROP#100;
nr=nrow(prop);
cumpct=j(nr,nr,0);
do i=1 to nr;
 cumpct(i,1:i)=1;
 end;
cumpct=cumpct * prop;
EIGEN=EIGEN||PROP||cumpct; ENAME='EIGENVAL' 'PERCENT' 'CUMPCT';
eigen=eigen(1:maxc,);
OUTPUT EIGEN DATA=EIGEN(RENAME=(ROW=COMP)) ROWNAME=COMPNAME COLNAME=ENAME;
FREE EIGEN PROP CUMPCT NR ENAME;
COMPNAME=%STR(COMPNAME%');

%IF &CORRECT=YES %THEN %DO;
 SCALE=S(1,)#/SQRT(S(1,)+1);
 RCOORD=DIAG(1#/SQRT(R)) * U * SQRT(DIAG(S+1)); RCOORD=RCOORD*SCALE;
 CCOORD=DIAG(1#/SQRT(C)) * V * SQRT(DIAG(S+1)); CCOORD=CCOORD*SCALE;
 %END;
%ELSE %DO;
 RCOORD=DIAG(1#/SQRT(R)) * U * DIAG(S);
 CCOORD=DIAG(1#/SQRT(C)) * V * DIAG(S);
 %END;
FREE U V;

%IF &BURT = YES AND &ROWRECON = YES %THEN %DO;
```

```
/* RECONSTRUCT ROW SCORES FROM COLUMN SCORES */
   FREE RCOORD;
   FETCH X DATA=_NOMISS_;
   RCOORD=(X*CCOORD);
   A=1#/X(,+); FREE X;
   DO I=1 TO NCOL(RCOORD);
     RCOORD(,I)=(A#RCOORD(,I))#/S(I,);
     END;
   FREE A;
   FETCH RNAME TYPE=CHAR DATA=_NOMISS_(KEEP=&ID);
%END;

%IF &CONTRIB=YES %THEN %DO;
 CABSCON = DIAG(C) * (CCOORD##2); FREE C;
 CABSCON = CABSCON * DIAG(1#/CABSCON(+,));
 CRELCON = CCOORD##2;
 CRELSUM = CRELCON(,+);
 DO I=1 TO NROW(CRELCON);
   CRELCON(I,)=CRELCON(I,)#/CRELSUM(I,1);
   END;
 FREE CRELSUM;
 cabscon=cabscon(,1:maxc);
 crelcon=crelcon(,1:maxc);
 OUTPUT CABSCON DATA=CABSCON(RENAME=(ROW=ID)) ROWNAME=CNAME COLNAME=COMPNAME;
 OUTPUT CRELCON DATA=CRELCON(RENAME=(ROW=ID)) ROWNAME=CNAME COLNAME=COMPNAME;
 FREE CABSCON CRELCON;

 %IF &BURT = NO %THEN %DO;
   RABSCON = DIAG(R) * (RCOORD##2); FREE R;
   RABSCON = RABSCON * DIAG(1#/RABSCON(+,));
   RRELCON = RCOORD##2;
   RRELSUM = RRELCON(,+);
   DO I=1 TO NROW(RRELCON);
     RRELCON(I,)=RRELCON(I,)#/RRELSUM(I,1);
     END;
   FREE RRELSUM;
   rabscon=rabscon(,1:maxc);
   rrelcon=rrelcon(,1:maxc);
   OUTPUT RABSCON DATA=RABSCON(RENAME=(ROW=ID)) ROWNAME=RNAME COLNAME=COMPNAME;
   OUTPUT RRELCON DATA=RRELCON(RENAME=(ROW=ID)) ROWNAME=RNAME COLNAME=COMPNAME;
   FREE RABSCON RRELCON;
   %END;
 %END;

%IF &BURT = NO OR &ROWRECON = YES %THEN %DO;
 rcoord=rcoord(,1:maxc);
 OUTPUT RCOORD OUT=RCOORD(RENAME=(ROW=ID)) ROWNAME=RNAME COLNAME=COMPNAME;
%END;
ccoord=ccoord(,1:maxc);
OUTPUT CCOORD OUT=CCOORD(RENAME=(ROW=ID)) ROWNAME=CNAME COLNAME=COMPNAME;

TITLE 'CORRESPONDENCE ANALYSIS';
PROC PRINT DATA=EIGEN NOOBS; VAR COMP EIGENVAL PERCENT CUMPCT;
 FORMAT EIGENVAL 8.5 PERCENT CUMPCT 8.2;
 TITLE3 'Inertias (eigenvalues) and proportion of total inertia';

%IF &BURT = NO %THEN %DO;
PROC PRINT NOOBS DATA=RCOORD;
 TITLE3 'Coordinates (optimal scale values) for rows';
%END;
```

```
PROC PRINT NOOBS DATA=CCOORD;
 TITLE3 'Coordinates (optimal scale values) for columns';

%IF &CONTRIB = YES %THEN %DO;
%IF &BURT = NO %THEN %DO;
PROC PRINT NOOBS DATA=RABSCON;
 TITLE3 'Absolute contributions of row points to inertia of each axis';
 TITLE4 '(contributions sum to 1 within a component)';
%END;

PROC PRINT NOOBS DATA=CABSCON;
 TITLE3 'Absolute contributions of column points to inertia of each axis';
 TITLE4 '(contributions sum to 1 within a component)';

%IF &BURT = NO %THEN %DO;
PROC PRINT NOOBS DATA=RRELCON;
 TITLE3 'Relative contributions of row points to inertia of each axis';
 TITLE4 '(contributions sum to 1 across components)';
 TITLE5 '(i.e., squared column-axis correlations)';
%END;

PROC PRINT NOOBS DATA=CRELCON;
 TITLE3 'Relative contributions of column points to inertia of each axis';
 TITLE4 '(contributions sum to 1 across components)';
 TITLE5 '(i.e., squared column-axis correlations)';
%END;

DATA PLOTDATA;
 SET CCOORD
%IF &BURT = NO %THEN %DO;
     RCOORD
     %END;
  ;
PROC PLOT;
 PLOT COMP2*COMP1=ID / VREF=0 HREF=0;
 PLOT COMP3*COMP1=ID / VREF=0 HREF=0;
TITLE2 'PLOT OF ROW AND COLUMN POINTS';
%OUT:
%MEND;
OPTIONS SOURCE;
CMS FILEDEF DATAIN DISK EXAMPLE DATA A;
data cadata;
 INFILE DATAIN;
 INPUT NAME $ C2 C6 C8 C12 C13 C16 C17 C33;
%CORRESP (DATA=CADATA,ID=NAME,VAR=C2--C33,BURT=NO,N=5);
```

Appendix 2 - Morphological data

Morphological characters : T/L, A1, A2, P/L, W/L, D/L

k401	2.	4.	3.	5.	5.	5.	k402	2.	3.	3.	4.	3.	5.
k403	3.	6.	6.	4.	6.	3.	k404	3.	6.	6.	4.	5.	2.
k405	3.	5.	4.	6.	4.	3.	k406	3.	6.	6.	8.	7.	3.
k407	3.	4.	4.	5.	6.	4.	k408	2.	4.	3.	5.	5.	4.
k409	3.	7.	6.	5.	7.	3.	k410	3.	7.	5.	8.	6.	3.
k411	2.	3.	3.	4.	3.	4.	k412	2.	5.	5.	5.	4.	2.
k413	2.	5.	3.	6.	6.	3.	k414	1.	5.	4.	5.	4.	4.
k415	2.	5.	4.	4.	4.	3.	k416	4.	5.	4.	5.	5.	4.
k417	3.	7.	6.	6.	7.	4.	k418	3.	6.	4.	5.	6.	4.
k419	3.	5.	3.	5.	5.	4.	k420	3.	5.	4.	4.	4.	3.
w101	2.	7.	7.	4.	6.	3.	w102	3.	3.	3.	4.	4.	3.
w103	2.	4.	3.	4.	4.	3.	w104	2.	4.	3.	5.	5.	3.
w105	3.	3.	3.	4.	4.	3.	w106	2.	5.	3.	6.	6.	3.
w107	2.	6.	5.	5.	4.	3.	w108	2.	2.	3.	3.	2.	5.
w109	2.	6.	5.	5.	6.	2.	w110	2.	6.	7.	4.	6.	2.
w111	2.	5.	4.	4.	6.	4.	w112	3.	4.	4.	4.	5.	3.
w113	3.	6.	5.	4.	4.	2.	w114	2.	6.	5.	4.	7.	3.
w115	1.	3.	4.	3.	4.	3.	w116	1.	4.	3.	4.	4.	2.
w117	3.	5.	3.	3.	4.	4.	w118	3.	5.	3.	3.	5.	5.
w119	4.	6.	5.	2.	4.	3.	w120	1.	5.	3.	5.	5.	3.
k301	2.	6.	4.	3.	7.	4.	k302	3.	4.	4.	3.	2.	4.
k303	2.	7.	7.	3.	8.	3.	k304	3.	9.	4.	5.	7.	3.
k305	3.	9.	8.	4.	7.	3.	k306	4.	6.	4.	4.	5.	3.
k307	3.	6.	4.	4.	5.	3.	k308	3.	7.	5.	5.	6.	2.
k309	3.	7.	5.	4.	6.	2.	k310	3.	7.	6.	6.	6.	2.
k311	2.	6.	5.	6.	6.	3.	k312	4.	7.	5.	4.	6.	2.
k313	1.	7.	5.	7.	7.	3.	k314	1.	9.	5.	8.	8.	2.
k315	1.	7.	4.	5.	7.	4.	k316	2.	4.	3.	4.	5.	4.
k317	2.	4.	3.	3.	4.	5.	k318	1.	7.	6.	3.	6.	3.
k319	4.	6.	4.	5.	6.	3.	k320	4.	7.	5.	5.	6.	3.
wc01	4.	7.	4.	6.	7.	4.	wc02	4.	6.	4.	6.	7.	4.
wc03	3.	9.	4.	5.	7.	1.	wc04	3.	6.	3.	5.	6.	2.
wc05	1.	7.	6.	3.	5.	1.	wc06	1.	5.	5.	6.	6.	4.
wc07	3.	5.	4.	5.	4.	3.	wc08	4.	2.	2.	2.	1.	6.
wc09	4.	7.	3.	6.	8.	2.	wc10	3.	5.	5.	4.	5.	5.
wc11	3.	5.	3.	6.	7.	6.	wc12	2.	6.	4.	3.	6.	4.
wc13	2.	4.	3.	4.	5.	4.	wc14	2.	6.	4.	6.	7.	4.
wc15	4.	6.	3.	6.	6.	3.	wc16	3.	4.	3.	5.	4.	3.

wc17 4. 5. 4. 5. 6. 3. wc18 4. 6. 3. 6. 6. 3.
wc19 3. 5. 3. 3. 5. 3. wc20 4. 7. 6. 6. 8. 3.
k101 4. 7. 5. 7. 6. 3. k102 2. 5. 3. 4. 4. 4.
k103 3. 5. 3. 4. 4. 3. k104 2. 6. 3. 4. 5. 3.
k105 3. 6. 3. 3. 4. 4. k106 3. 5. 3. 4. 4. 3.
k107 3. 5. 3. 4. 4. 4. k108 3. 5. 3. 4. 4. 4.
k109 3. 6. 3. 4. 4. 9. k110 3. 6. 3. 4. 5. 4.
k111 3. 7. 4. 5. 6. 4. k112 2. 5. 3. 4. 4. 4.
k113 3. 6. 3. 4. 5. 4. k114 2. 6. 3. 4. 5. 3.
k115 3. 5. 2. 4. 4. 4. k116 3. 4. 3. 3. 4. 6.
k117 4. 3. 3. 3. 3. 7. k118 4. 4. 3. 4. 2. 3.
k119 2. 5. 3. 5. 4. 4. k120 2. 4. 4. 4. 4. 5.
k201 2. 7. 4. 5. 7. 3. k202 2. 4. 3. 4. 4. 6.
k203 2. 4. 3. 4. 4. 5. k204 3. 6. 4. 7. 5. 2.
k205 3. 7. 8. 5. 5. 3. k206 2. 4. 3. 3. 6. 4.
k207 2. 5. 4. 5. 5. 6. k208 3. 5. 3. 4. 5. 5.
k209 3. 5. 3. 5. 6. 3. k210 4. 7. 4. 6. 6. 3.
k211 4. 9. 6. 6. 7. 2. k212 4. 7. 4. 6. 7. 2.
k213 3. 6. 3. 6. 5. 2. k214 4. 7. 4. 6. 7. 3.
k215 2. 4. 3. 4. 4. 4. k216 3. 7. 4. 6. 6. 3.
k217 1. 4. 2. 6. 6. 5. k218 4. 5. 3. 4. 5. 4.
k219 1. 4. 3. 5. 5. 4. k220 3. 7. 4. 5. 7. 3.
w201 3. 4. 3. 5. 4. 3. w202 3. 6. 4. 5. 5. 2.
w203 3. 6. 4. 4. 5. 2. w204 3. 6. 6. 5. 5. 3.
w205 4. 4. 3. 4. 4. 4. w206 2. 3. 3. 2. 3. 4.
w207 1. 5. 3. 4. 5. 3. w208 3. 5. 4. 4. 6. 5.
w209 4. 8. 8. 6. 6. 1. w210 3. 5. 5. 3. 5. 2.
ta01 5. 3. 2. 2. 2. 5. ta02 4. 3. 2. 3. 3. 5.
ta03 6. 3. 3. 2. 3. 4. ta04 4. 7. 7. 4. 6. 5.
si01 5. 3. 3. 2. 3. 7. si02 4. 5. 3. 6. 5. 3.
si03 5. 3. 2. 4. 4. 6. si04 3. 3. 2. 3. 3. 6.
si05 7. 3. 3. 2. 3. 4. si06 4. 5. 3. 3. 5. 4.
si07 5. 5. 4. 3. 4. 4. si08 2. 4. 3. 4. 4. 4.
si09 4. 7. 6. 6. 7. 2. si10 8. 5. 3. 3. 4. 4.
si11 8. 4. 2. 3. 3. 5. si12 3. 2. 2. 2. 2. 5.
si13 5. 1. 1. 2. 2. 6. si14 3. 4. 3. 5. 5. 4.
si15 6. 2. 2. 2. 3. 4. si16 5. 2. 2. 2. 2. 6.
si17 3. 4. 3. 3. 4. 3. si18 3. 5. 3. 4. 4. 4.
si19 3. 6. 3. 6. 7. 4. si20 5. 2. 2. 1. 2. 7.
p101 5. 2. 2. 1. 1. 4. p102 5. 2. 2. 2. 2. 4.
p105 7. 3. 3. 2. 4. 7. p106 6. 1. 2. 1. 1. 5.
p107 5. 2. 2. 1. 1. 5. p108 5. 2. 3. 2. 2. 5.
p109 6. 2. 2. 2. 2. 5. p110 5. 2. 2. 1. 2. 5.

pl11	5.	1.	2.	1.	1.	6.	pl12	5.	1.	1.	1.	1.	5.
cr01	5.	5.	3.	3.	5.	4.	cr02	4.	5.	3.	4.	3.	4.
cr03	4.	4.	3.	3.	4.	4.	cr04	5.	6.	3.	5.	6.	3.
cr05	4.	6.	4.	4.	6.	3.	cr06	2.	3.	2.	3.	3.	5.
cr07	2.	4.	3.	2.	3.	4.	cr08	7.	2.	2.	1.	2.	5.
cr09	4.	7.	5.	6.	8.	3.	cr10	3.	3.	2.	3.	3.	5.
cr11	3.	5.	4.	2.	3.	3.	cr12	2.	2.	2.	2.	2.	6.
cr13	5.	4.	3.	3.	5.	6.	cr14	4.	3.	2.	3.	3.	5.
cr15	4.	5.	3.	5.	6.	4.	cr16	4.	2.	2.	3.	3.	6.
cr17	6.	4.	3.	4.	5.	4.	cr18	6.	4.	4.	3.	3.	3.
cr19	6.	5.	4.	3.	4.	3.	cr20	3.	7.	4.	5.	7.	2.
dm01	9.	5.	5.	2.	4.	3.	dm03	6.	3.	4.	1.	3.	6.
dm04	3.	7.	8.	5.	5.	1.	dm05	3.	6.	5.	5.	7.	3.
dm06	6.	3.	2.	2.	4.	6.	dm07	5.	2.	3.	2.	2.	5.
dm08	4.	7.	5.	6.	7.	2.	dm09	6.	3.	3.	1.	2.	6.
dm10	4.	7.	5.	6.	6.	2.	dm11	4.	7.	4.	6.	7.	3.
dm12	4.	7.	3.	5.	6.	1.	dm13	3.	6.	3.	4.	5.	3.
dm14	3.	7.	5.	3.	5.	2.	dm15	4.	2.	2.	2.	1.	6.
dm16	3.	6.	5.	4.	5.	3.	dm17	3.	6.	5.	5.	5.	2.
dm18	6.	3.	3.	1.	2.	4.	dm19	5.	3.	2.	2.	3.	5.
dm20	6.	4.	3.	3.	3.	5.	po01	4.	9.	8.	7.	7.	2.
po02	5.	7.	6.	8.	4.	2.	po03	5.	7.	3.	4.	4.	2.
po04	2.	3.	2.	4.	4.	5.	po05	3.	7.	4.	6.	5.	4.
po06	3.	7.	4.	6.	5.	4.	po07	4.	6.	4.	6.	5.	6.
po08	3.	5.	3.	4.	5.	5.	po09	3.	5.	3.	4.	5.	4.
po10	3.	5.	4.	4.	4.	4.	po11	3.	5.	3.	4.	5.	4.
po12	3.	7.	5.	6.	5.	3.	po13	3.	5.	4.	5.	5.	4.
po14	3.	7.	4.	6.	5.	3.	po15	3.	7.	5.	6.	5.	3.
po16	3.	7.	4.	6.	5.	3.	po17	3.	7.	4.	6.	5.	3.
po18	3.	7.	3.	6.	5.	3.	bl01	3.	9.	8.	6.	6.	1.
bl02	2.	7.	6.	5.	6.	2.	bl03	2.	7.	5.	5.	5.	5.
bl04	2.	9.	5.	6.	8.	2.	bl05	2.	7.	5.	5.	6.	2.
bl06	3.	7.	7.	5.	5.	2.	bl07	3.	9.	8.	6.	6.	1.
bl08	3.	7.	7.	6.	6.	2.	bl09	3.	7.	5.	4.	6.	3.
bl10	2.	6.	5.	5.	5.	3.	bl11	2.	7.	5.	5.	6.	2.
bl12	2.	6.	5.	4.	5.	3.	bl13	5.	7.	4.	7.	7.	4.
bl14	2.	7.	4.	5.	6.	3.	bl15	3.	7.	6.	4.	5.	3.
bl16	3.	6.	4.	4.	5.	3.	bl17	3.	7.	6.	3.	8.	2.
bl18	3.	5.	3.	5.	4.	3.	bl19	3.	5.	3.	4.	4.	3.
bl20	3.	5.	2.	4.	4.	3.	wf01	3.	3.	2.	4.	5.	5.
wf02	4.	2.	2.	2.	2.	6.	wf03	5.	2.	2.	2.	3.	6.
wf04	3.	5.	4.	4.	3.	4.	wf05	5.	1.	1.	2.	1.	6.
wf06	4.	2.	2.	2.	1.	5.	wf07	5.	3.	4.	2.	3.	5.

wf08 5. 2. 2. 2. 2. 6. wf09 6. 3. 3. 3. 4. 6.
wf10 5. 2. 2. 2. 3. 5. wf11 5. 2. 2. 3. 3. 6.
wf12 4. 9. 8. 6. 8. 2. wf13 4. 6. 4. 5. 8. 4.
wf14 3. 7. 6. 5. 5. 2. wf15 3. 6. 3. 6. 5. 3.
wf16 6. 4. 3. 2. 4. 5. wf17 6. 4. 4. 2. 4. 6.
wf18 3. 7. 4. 6. 6. 3. wf19 3. 5. 3. 4. 5. 4.
wf20 5. 7. 6. 5. 6. 2. ma06 5. 2. 2. 1. 3. 8.
ma07 6. 3. 3. 2. 4. 6. ma08 6. 3. 2. 3. 4. 5.
ma09 6. 2. 2. 2. 3. 5. ma10 6. 3. 2. 3. 4. 7.
wr01 4. 5. 3. 4. 6. 4. wr02 4. 5. 3. 4. 6. 5.
wr03 4. 5. 3. 4. 6. 5. wr04 2. 9. 3. 6. 7. 2.
wr05 6. 2. 2. 1. 2. 7. wr06 3. 2. 2. 5. 5. 3.
wr07 3. 7. 3. 4. 7. 3. wr08 3. 7. 5. 4. 5. 2.
wr09 4. 3. 2. 3. 3. 5. wr10 4. 3. 3. 2. 3. 5.
wr11 6. 4. 3. 4. 4. 4. wr12 4. 4. 2. 3. 4. 4.
wr13 4. 3. 3. 2. 4. 6. wr14 5. 3. 2. 2. 3. 5.
wr15 6. 2. 2. 2. 2. 5. wr16 4. 6. 3. 5. 8. 4.
wr17 6. 2. 2. 2. 2. 6. wr18 7. 2. 2. 2. 3. 5.
wr19 7. 3. 3. 2. 3. 5. wr20 3. 7. 6. 6. 6. 3.
c101 3. 6. 3. 5. 7. 4. c102 3. 5. 3. 5. 4. 2.
c103 3. 5. 3. 4. 6. 4. c104 4. 7. 4. 6. 6. 2.
c105 3. 6. 3. 6. 6. 3. c106 3. 4. 3. 4. 5. 3.
c107 3. 3. 2. 4. 4. 5. c108 3. 7. 7. 5. 5. 2.
c109 3. 6. 3. 5. 6. 4. c110 4. 1. 1. 1. 1. 6.
c111 5. 4. 3. 4. 4. 3. c112 5. 9. 6. 6. 7. 1.
c113 4. 5. 3. 4. 6. 3. c114 5. 7. 4. 5. 6. 2.
c115 6. 6. 3. 5. 6. 3. c116 4. 5. 3. 5. 6. 4.
c117 4. 5. 3. 6. 6. 4. c118 3. 6. 3. 4. 5. 3.
c119 4. 7. 3. 7. 5. 4. c120 2. 4. 3. 6. 4. 3.
so01 6. 3. 3. 2. 3. 5. so02 5. 3. 3. 2. 3. 5.
so03 5. 3. 3. 2. 2. 4. so04 8. 2. 2. 1. 2. 5.
so05 6. 3. 3. 2. 2. 5. so06 5. 2. 2. 2. 2. 4.
so07 6. 3. 3. 3. 3. 6. so08 5. 3. 3. 3. 2. 5.
so09 7. 3. 3. 3. 3. 3. so10 5. 4. 6. 2. 4. 4.
so11 8. 2. 3. 2. 2. 6. so12 6. 2. 3. 2. 2. 4.
so13 6. 3. 3. 2. 2. 4. so14 8. 4. 4. 2. 3. 4.
so15 6. 6. 3. 4. 4. 2. so16 6. 2. 3. 1. 2. 5.
so17 2. 3. 3. 2. 2. 6. so18 7. 4. 4. 3. 3. 4.
so19 6. 3. 3. 3. 4. 5. so20 6. 2. 2. 2. 2. 5.
hr01 6. 9. 7. 6. 7. 2. hr02 4. 5. 3. 4. 5. 2.
hr03 4. 7. 6. 4. 5. 1. hr04 4. 6. 4. 5. 5. 2.
hr05 5. 9. 9. 5. 6. 1. hr06 3. 7. 6. 5. 5. 2.
hr07 4. 9. 5. 5. 6. 1. hr08 2. 7. 4. 6. 6. 3.

hr09	2.	5.	4.	6.	5.	3.	hr10	2.	5.	3.	6.	6.	3.
hr11	2.	4.	3.	4.	4.	3.	hr12	3.	4.	3.	5.	5.	4.
hr13	5.	9.	4.	5.	7.	1.	hr14	2.	5.	3.	6.	5.	3.
hr15	4.	7.	3.	5.	6.	2.	hr16	4.	7.	4.	7.	7.	1.
hr17	3.	6.	4.	5.	6.	2.	hr18	4.	7.	5.	6.	7.	1.
hr19	2.	5.	4.	4.	4.	2.	hr20	2.	5.	4.	4.	4.	1.
fc01	6.	7.	6.	4.	8.	3.	fc02	4.	5.	3.	5.	5.	3.
fc03	4.	6.	4.	7.	5.	2.	fc04	4.	6.	4.	4.	6.	3.
fc05	4.	6.	4.	5.	6.	4.	fc06	3.	5.	4.	3.	4.	3.
fc07	3.	5.	3.	5.	5.	3.	fc08	3.	5.	4.	5.	5.	3.
fc09	2.	4.	4.	4.	3.	3.	fc10	3.	4.	3.	6.	5.	3.
fc11	3.	4.	3.	5.	4.	3.	fc12	4.	4.	3.	6.	4.	3.
fc13	4.	5.	5.	6.	6.	7.	fc14	4.	5.	3.	6.	5.	4.
fc15	3.	3.	3.	4.	4.	4.	fc16	3.	5.	3.	5.	5.	3.
fc17	5.	6.	3.	8.	5.	2.	fc18	3.	5.	5.	4.	5.	4.
fc19	2.	4.	3.	4.	3.	3.	fc20	3.	5.	4.	5.	5.	3.

Appendix 3 - flavonoid data

flavonoid compound # : 1, 2, 3, 5, 6, 12, 13, 16, 20, 21, 22, 23, 24, 25, 33, 37, 26, 17

CR01	111111001101111001	CR02	111110011111101001
CR03	111111001101111001	CR04	111111000101110000
CR05	111111000101110000	CR06	111111001101111000
CR07	111110011111100001	CR08	111110011111101001
CR10	111111011101111001	CR09	111111101101111001
CR11	111010011111101001	CR12	111111010101111000
CR13	111110010111101000	CR14	111111000101111000
CR15	111111000101111000	CR16	111011001101111001
CR17	111110011111101001	CR18	111110011111101000
CR19	111110011111101000	CR20	111111101101110001
K101	111111101101111100	K102	111111100101111000
K103	111111100101111000	K104	111111101101111000
K105	111111101101111100	K106	111111100101110000
K107	111111100101110000	K108	111111100101110000
K109	111111101101111001	K110	111111101101111000
K111	111111100101110000	K112	111111100101110000
K113	111111100101111000	K114	111111101101111101
K115	111111100101110000	K116	111111101101110000
K117	111111000101110000	K118	111111101101111001
K119	111111000101110000	K120	111111000101110001
PL01	111110010111101000	PL02	111110010111101000
PL05	111010011111101001	PL06	111110011111101001
PL07	111110011111101001	PL08	111110011111101000
PL09	111010011111101001	PL10	111110011111101001
PL11	111110011111101001	PL12	111110011111101001
TA01	111110011111101001	TA02	111110010111101000
TA03	111011001101111001	TA04	111111101101111000
DM01	111010011111101001	DM03	111110011111101001
DM04	111011000101111000	DM05	111011100101110001
DM06	111110011111101001	DM07	111010011111101001
DM08	111111100101110001	DM09	111010011111101001
DM10	111101001101110000	DM11	111001101101111000
DM12	111001001101111000	DM13	111011100101110001
DM14	111111000101110000	DM15	111010011111101000
DM16	111011000101110000	DM17	111011000101110000
DM18	111110011111101001	DM19	111110011111101001
DM20	111010011111101001	PO01	111111101101110001
PO02	111111101101111001	PO03	111111101101110101

PO04 1111111011011110000 PO05 1110010011011110011
PO06 1110010011011110010 PO07 1110011011011110011
PO08 1110010011011110010 PO09 1110011011011110010
PO10 1110011011011110010 PO11 1110011011011110011
PO12 1110011011011110011 PO13 1110010011011110010
PO14 1110010011011110011 PO15 1110011011011110011
PO16 1110011011011110011 PO17 1110011011011110011
PO18 1110111011011110001 W101 1111110011011111000
W102 1111111011011111001 W103 1111111001011111001
W104 1111111011011111101 W105 1110110011011111000
W106 1111111011011110001 W107 1111110011011110000
W108 1111111011011111000 W109 1111110011011111001
W110 1110011011011111011 W111 1110010011011110011
W112 1110010011011110011 W113 1111110011011111000
W114 1110110011011110001 W115 1111110001011110001
W116 1110110011011110000 W117 1110011011011110011
W118 1111111011011110001 W119 1111111011011110001
W120 1111111011011111001 BL01 1111110011011110001
BL02 1110110011011110001 BL03 1110110011011110001
BL04 1111111011011110100 BL05 1111111011011111001
BL06 1111110011011110001 BL07 1111111011011111001
BL08 1111111011011111001 BL09 1111111011011111001
BL10 1110011011011110001 BL11 1110010011011110000
BL12 1111111011011111001 BL13 1110011011011110000
BL14 1111011011011110000 BL15 1111111001011110001
BL16 1111111001011110001 BL17 1111110011011111001
BL18 1111111011011110100 BL19 1111111011011110101
SI01 1111100111111101001 SI02 1111111011011111101
SI03 1111110111011111001 SI04 1111111011011111000
SI05 1111100111111101001 SI06 1111111001011111000
SI07 1111111011011111000 SI08 1111111011011110101
SI09 1111111011011110101 SI10 1111100111111101001
SI11 1111100111111Ô1001 SI12 1111111011011111001
SI13 1111100111111101001 SI14 1111111011011111100
SI15 1111100111111101001 SI16 1111100111111101000
SI17 1111110001011110000 SI18 1111111001011110000
SI19 1111111011011110000 SI20 1111100111111101001
CL01 1111111001011110001 CL02 1111111001011110001
CL03 1111111101011111000 CL04 1111111101011111001
CL05 1111111101011111001 CL06 1111111001011110101
CL07 1111110001011111000 CL08 1111110001011111000
CL09 1110111001011111001 CL10 1111100101111101000
C111 1111110001011110000 CL12 1111111001011111001

CL13 1111110001011100000 CL14 1111111001011100000
CL15 1111111001011100001 CL16 1110110001011100000
CL17 1111110001011100000 CL18 1111111001011100001
CL19 1111111011011110000 CL20 1111111001011100000
HR01 1111111011011110001 HR02 1111111011011100001
HR03 1110111011011100011 HR04 1111111001011110111
HR05 1111110001011110001 HR06 1110111001011110000
HR07 1111110001011110001 HR08 1111111011011110000
HR09 1111111011011110001 HR10 1110110001011110000
HR11 1110110001011100000 HR12 1110110001011100000
HR13 1110111001011110001 HR14 1111111011011100000
HR15 1111111011011111100 HR16 1111111001011100000
HR17 1111111011011110000 HR18 1110111001011111101
HR19 1111111001011110001 HR20 1111111011011110001
WR01 1111111011011111101 WR02 1111111011011111101
WR03 1111111011011111101 WR04 1111110011011100001
WR05 1111100111111101001 WR06 1111111011011111100
WR07 1111111011011100001 WR08 1110111001011100001
WR09 1111111111011110000 WR10 1111100001011110000
WR11 1110111111011110001 WR12 1111111011011110001
WR13 1111111011011110001 WR14 1111111011011110001
WR15 1110100111111101001 WR16 1111111011011111101
WR17 1110100111111101001 WR18 1110100111111101001
WR19 1110100111111101001 WR20 1111111001011110000
FC01 1111111001011100000 FC02 1111110001011100000
FC03 1111111001011111101 FC04 1110111001011100001
FC05 1111110011011100000 FC06 1111111001011111101
FC07 1110111001011100000 FC08 1110111001011101101
FC09 1110111001011100001 FC10 1111111011011110000
FC11 1111111011011100000 FC12 1110111001011110000
FC13 1111110001011110000 FC14 1111111001011110000
FC15 1110111001010100000 FC16 1111111101011110000
FC17 1111111001011111100 FC18 1111110001011100000
FC19 1111111001011100001 FC20 1111111001011101101
WF01 1111111011011110001 WF02 1111100111111101001
WF03 1111100111111101001 WF04 1111111011011110001
WF05 1110100111111101001 WF06 1110100111111101001
WF07 1111100111111101001 WF08 1111100111111101001
WF09 1111100111111101001 WF10 1110100111111101001
WF11 1110100111111101001 WF12 1111111001011100000
WF13 1111111001011110001 WF14 1110111011011100001
WF15 1111111011011100000 WF16 1111100111111101001
WF17 1111100111111101001 WF18 1111111011011100001

WF20	1111110111011111001	SO01	1111100111111101001
SO02	1111100111111101001	SO03	1110100111111101001
SO04	1110100111111101001	SO05	1110100111111101001
SO06	1111100111111101001	SO07	1111100111111101001
SO08	1110100111111101001	SO09	1111100111111101001
SO10	1111100111111101001	SO11	1110100111111101001
SO12	1111100111111101001	SO13	1111100111111101001
SO14	1110100111111101001	SO15	1111100111111101000
SO16	1111100111111101001	SO17	1111110111011111001
SO18	1110100111111101001	SO19	1111100111111101001
SO20	1111100111111101001	K301	1111111001011101111
K302	1111111001011100111	K303	1111111001011100001
K304	1111111011011110000	K305	1111100001011110000
K306	1111111011011100001	K307	1111110001011100000
K308	1111100001011100000	K309	1111100001011100000
K310	1111100001011100000	K311	1111111011011100001
K312	1111100001011100000	K313	1111111001011100001
K314	1110111001011100001	K315	1111111011011110001
K316	1110110001011100111	K317	1110010011011100101
K318	1111110001011110001	K319	1110110001011100101
K320	1111110011011100000	WC01	1111111011011111101
WC02	1111111011011101101	WC03	1111110001011100000
WC04	1110111011011110000	WC05	1111111011011100000
WC06	1111111011011111101	WC07	1111111011011101011
WC08	1111110001011100000	WC08	1111110111011110000
WC10	1111111001011111101	WC11	1110010011011100101
WC12	1111111011011101011	WC13	1111111001011101001
WC14	1110111011011101001	WC15	1110111011011101001
WC16	1110110011011110011	WC17	1110111001011100011
WC18	1111111011011111001	WC19	1111111011011110011
WC20	1111111011011110000	W201	1111110011011100011
W202	1111111011011100011	W203	1111111011011100011
W204	1111111011011100011	W205	1111110011011100000
W206	1111111011011100011	W207	1111110011011100011
W208	1111110011011100011	W209	1111110011011110011
W210	1111111011011100011	K401	1111111001011111101
K402	1111111001011101001	K403	1111111001011111001
K404	1111111001011111001	K405	1111110001011100000
K406	1111111011011111101	K407	1111111001011101001
K408	1111111001011111001	K409	1110111001011100000
K410	1111111011011100000	K411	1110111001011101011
K412	1111111011011100011	K413	1111111001011100101
K414	1110111001011110000	K415	1111110001011100000

K416 1111111001011110000 K417 1111111001011100000
K418 1110111001011110000 K419 1110011011011100011
K420 1111111011011110001 K201 1111100001011110000
K202 1111011011011100010 K203 1111111011011100011
K204 1111111001011110001 K205 1111111001011110001
K206 1111111001011100001 K207 1111110001011110000
K208 1111110001011100000 K209 1111111001011110001
K212 1111111001011110001 K213 1111111001011100000
K214 1110111001011100001 K215 1111111001011110000
K216 1111111011011110001 K217 1111111001011100001
K218 1111111011011100001 K219 1111111011011100001
K220 1111111001011100001

Appendix 4 - Electrophoresis data

Enzyme loci : AAT1, AAT2, ADH1, CE1, CE2, CE3, IDH1, IDH2, LAP1, LAP2, PGI1 ,PGI2, PGM1, SKDH1, MDH1, MDH2, MDH4

SI01	11111112111331131	SI02	11113112111331132
SI03	13112111111331132	SI04	13112111111331132
SI05	13111111111211132	SI06	13111111111321231
SI07	13113111111331132	SI08	11111111111331232
SI09	11113111111331232	SI10	11111111111131132
SI11	13111112111331132	SI12	13111111111331132
SI13	13111111111121132	SI14	11112112111311132
SI15	12111111111131133	SI16	11111111111221131
SI17	11112112111331132	SI18	11113112111331132
SI19	13112111111321232	SI20	13111111111211132
TA01	13111111111211132	TA02	11111111111211132
TA04	13111112111321232	PL01	11111111111231132
PL02	11111111111231132	PL05	11111111111231132
PL06	11111111111231132	PL07	11111111111231132
PL08	11111111111231132	PL09	11111111111231132
PL10	11111111111231132	PL11	11111111111231132
PL12	11111111111231132	MA06	13111111111211132
MA07	13111111111211132	MA08	13111111111211132
MA09	13111111111211132	MA10	13111111111211132
W201	11111111111331132	W202	13114111111331132
W203	13114111111331132	W204	13114111111331132
W205	13114111111331132	W206	11111111111331132
W207	11114111111311132	W208	11111111111331132
W209	11112111111311132	W210	11113111111331132
WR01	11111112111331132	WR02	11111112111331132
WR03	11111112111331132	WR04	13114112111321132
WR05	11111111111311133	WR06	11113112111311132
WR07	13111112111321132	WR08	11112111111331132
WR09	11111112111321132	WR10	11113111111331233
WR11	11111112111321132	WR12	12114112111331132
WR13	11113112111331233	WR14	12114112111331112
WR15	11111112111231112	WR16	13111112111321232
WR17	11111112111231112	WR18	11111111111231133
WR19	11111111111231133	WR20	13114112111311132
FC01	11114121111321234	FC02	13113112111331132
FC03	11114112111311132	FC04	12112111111331232
FC05	11114111111311132	FC06	11112111111331232
FC07	11114112111331132	FC08	13113112111331132

FC09	11114111111331132	FC10	11114111111311132
FC11	11112112111331232	FC12	12112111111331132
FC13	12114111111331132	FC14	12112112111331132
FC15	11112111111331132	FC16	13112111111311132
FC17	11112111111231134	FC18	13114111111231132
FC19	13114111111231132	FC20	13112111111331132
SO01	11113112111311134	SO02	11111112111211132
SO03	11113111111211132	SO04	11114112111331232
SO05	11113112111331134	SO06	11111111111331132
SO07	13113112111211132	SO08	11113111111231132
SO09	11111111111211132	SO10	13111112111211132
SO11	13111111111331132	SO12	11113112111321134
SO13	12113112111231132	SO14	11113111111211134
SO15	13113112111311132	SO16	13113112111311132
SO17	13113111111211132	SO18	11113111111211132
SO19	13111111111221132	SO20	13114111111221132
CL01	13113111111311132	CL02	11113111111331232
CL03	11112111111331232	CL04	11113112111311132
CL05	11114112111331232	CL06	11113112111331132
CL07	11112112111331232	CL08	11111112111311132
CL09	12113111111311132	CL10	13114111111321132
CL11	11113112111311234	CL12	11111111111331232
CL13	11113111111331134	CL14	11113112111311132
CL15	13113111111331132	CL16	11114111111331232
CL17	11114111111331232	CL18	11111112111321232
CL19	11114111111311132	CL20	11113111111321222
CR01	13114112111331132	CR02	13114111111321132
CR03	13114111111331132	CR04	13111112111111232
CR05	13111112111111232	CR06	13111112111331132
CR07	13111112111331132	CR08	11111112111231132
CR09	11113112111131132	CR10	13111112111331132
CR11	12114112111331132	CR12	13111112111331132
CR13	13114112111331132	CR14	13113112111311132
CR15	11113112111331132	CR16	13114112111331132
CR17	13114112111331132	CR18	13114112111331132
CR19	13114111111331132	CR20	11111111111121132
K401	11111111111311132	K402	11112111111211132
K403	11113112111211234	K404	11113112111311234
K405	11112111111311232	K406	11112111111311132
K407	11113112111221132	K408	11114112111211132
K409	11114112111221232	K410	11112111111221132
K411	11111112111211232	K412	11114111111231132
K413	11111111111221132	K414	11111111111211132

K415	111141121113111132	K416	111121111113211132
K417	111111121112311132	K418	111111111112111132
K419	111131111112111132	K420	111141111112311132
K201	111131111113312232	K202	111111111113312232
K203	111111111113312232	K204	111131111113111132
K205	111141111113111132	K206	111131111113211132
K207	111141111113311132	K208	111141121113212232
K209	111131111113211132	K210	111131111113311132
K211	111131111113312232	K212	111131111113112232
K213	111141111113311132	K214	111131111113312232
K215	111131111113312232	K216	111141111113111132
K217	111131111113311132	K218	111111111113211132
K219	111141111113211132	K220	111141111113111132
K101	111141111113111134	K102	111131111113111134
K103	111131111113111134	K104	111131111113111134
K105	111131111113111134	K106	111131111113111134
K107	111131111113111134	K108	111131111113111134
K109	111131111113111134	K110	111131111113111134
K111	111131111113111134	K112	111131111113111134
K113	111131111113111134	K114	111131111113111134
K115	111131111113111134	K116	111121111113111134
K117	111141111113111134	K118	111121111113111134
K119	111121111113311134	K120	111121111113111134
BL01	111141111113311132	BL02	111141111113311132
BL03	111141111113311132	BL04	111121121113111132
BL05	111141111113311132	BL06	111141111113311132
BL07	111141111113311132	BL08	111141111113311132
BL09	111121121113311132	BL10	111141111113311132
BL11	111141111113311132	BL12	111141111113311132
BL13	111131111113111132	BL14	111121121113311132
BL15	111131111113111132	BL16	111131111113111132
BL18	111121111113311132	BL19	111121111113311132
BL20	111121111113311132	K301	111121111113111134
K302	131131111113112234	K303	111121111113111134
K304	131111111113112234	K305	131121121113312234
K306	131111121113112234	K307	111131111113312234
K308	131121111113312234	K309	131121111113312234
K310	131121111113312234	K311	111121111113112234
K312	131121111113312234	K313	111121111113112234
K314	111121111113112234	K315	111121111113112234
K316	121131121113112234	K317	111131111113112234
K318	111131111113312234	K319	111111121113112234
K320	111131121113112234	W101	111121111113211112

W102	13111111111311112	W103	11114111111321112
W104	11114112111321112	W105	13111112111311112
W106	11114111111331112	W107	12111111111331112
W108	13114111111331134	W109	11114111111321132
W110	13114111111321132	W111	11112111111321112
W112	11114111111311112	W113	13113111111331132
W114	12111111111331112	W115	11111111111331134
W116	11113111111331132	W117	11113111111311234
W118	11113111111311234	W119	11113111111311234
W120	11113111111331134	DM01	12111111111311131
DM03	12114112111321132	DM04	13112112111311232
DM05	13111111112331222	DM06	12111112111331233
DM07	11111111111211232	DM08	12113112111331234
DM09	13111111111321232	DM10	13112111111321232
DM11	13113111112331132	DM12	12113112112331132
DM13	11113111112311132	DM14	11111112111311132
DM15	12111111111231232	DM16	13112111112321214
DM17	13112111112321114	DM18	13111111111311112
DM19	13111111111231112	DM20	12111111111231112
WF01	11111111111311138	WF02	11111112111211133
WF03	13111111111211133	WF04	13114111111331233
WF05	13111111111311133	WF06	11111111111211133
WF07	13114111111231133	WF08	13111111111331133
WF09	13114112111211133	WF10	11114112111211133
WF11	13114111111311133	WF12	11112111111311233
WF13	13113112111331133	WF14	11112111111331133
WF15	11113111111311133	WF16	11111111111211133
WF17	11111111111211133	WF18	11113111111311133
WF19	13112111111311133	WF20	11113111111331133
WC01	11114111111331132	WC02	13112111111311132
WC03	12112111111331232	WC04	13112111111331132
WC05	11112112111331232	WC06	13111111111331132
WC07	11114111111331132	WC08	11111111111331232
WC09	11111111111331132	WC10	13111111111331232
WC11	11111111111311234	WC12	11114111111331234
WC13	13114111111321132	WC14	11114111111331132
WC15	11114112111311234	WC16	13114112111331134
WC17	11114111111331134	WC18	13114112111331132
WC19	13114112111331132	WC20	13111112111331132
PO01	12114111111311132	PO02	12114111111311132
PO03	11112111111311132	PO04	11114111111311132
PO05	11113111111331132	PO06	12113111111331132
PO07	12113111111331132	PO08	11113111111311132

PO09 11113111111311132 PO10 11113111111311132
PO11 12113111111311132 PO12 12113111111311132
PO13 11113111111311132 PO14 13113111111331132
PO15 13113111111331132 PO16 12113111111311132
PO17 12113111111331132 PO18 13113111111331132
HR01 12112111111321134 HR02 11114112111311134
HR03 12114122111331234 HR04 12114112111331232
HR05 11114111111331132 HR06 11111112111321234
HR07 13112111111311134 HR08 11112111111341134
HR09 13113111111351234 HR10 11114111111331234
HR11 11112111111331232 HR12 11112111111331232
HR13 12114111111331134 HR14 12111112111321232
HR15 11114112111331132 HR16 12114112111311234
HR17 11111111111311134 HR18 11113111111321132
HR19 11114112111311132 HR20 11114112111311132

Transformation key

AAT1 : AA=1
AAT2 : AA=1, BB=2, AB=3
ADH1 : AA=1
CE1 : AA=1
CE2 : AA=1, BB=2, AB=3, nil=4
CE3 : AA=1
IDH1 : AA=1, AB=2
IDH2 : AA=1, BB=2
LAP1 : AA=1
LAP2 : AA=1
MDH1 : AA=1, AB=2
MDH2 : AA=1, BB=2, AB=3
MDH4 : nil=1, AA=2, BB=3, AB=4
PGI1 : AA=1, BB=2
PGI2 : AA=1, BB=2, AB=3
PGM1 : AA=1, BB=2, AB=3, AC=4, AD=5
SKDH1 : AA=1